

167 ysAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrp 18


```

670 ThrSerAsnMetHisGlySerLeuGlnGlnLysHisArgIleLeuH1 686
      |||
2433 .....GGAGCAGTGTCCAG.....ATTCATC 2455
686 slyLeuGlnAsnGlyAsnSerProAlaIleAlaLysIleThr. 702
      |||
2456 TACTATAGCCCAAAACAGCAGCCAGAGTGTGACGATACATTCATCTC 2505
703 .....AlaGlnAlaThrGlyLysAsp 709
2506 AGCAGCAGCAGATAGATTTTCTCAAGTCACACACTGTGACCAAAATTA 2555
710 ThrSerSerIleThrSerCysGlyAsp..... 718
2556 GTGACGTCTCTGTAGCTGTGTGGGAGTCATGTAAGTACTATGCT 2605
719 .....GlyAsnValAla.....LysGlnGlnG 726
2606 TATGGGCGACAGTGTGACGTCATATCTTCTTGTACACAAACAGCAAC 2655
726 LnuSerProLysLysGlyAsnAsnAlaLeuLeuArgTyrLeuLeu 742
2656 AG..... 2657
743 AspArgAspAspProSerAspAlaLeuSerLysGlnLeuGlnProGlnVa 759
2658 .....TCACAGACATTTTACAGACGACGACGACGACGACGA 2692
759 LglGlyValAspAsnLysMetSerGlnCysThrSerThrIleProS 776
2693 GCAG.....AGCTCCAGAGAGCAGCAGCTCAGTTCAGACAAACAT 2736
776 eiserSerGlnGlyLysAspProLysIleLysThrGlnThrSerGlnGlu 792
2737 CTCAGGCTCAGCTGACCCAGCA..... 2759
793 GlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLeuThrse 809
2759 ..... 2759
809 rSerAspPheTyrAsnAsnSerIleSerSerAsnGlySerHisLeuGly 826
2759 ..... 2759
826 hrlYsglnGlnValPheGlnGlyThrAsnSerLeu...GlyLeuLysSer 841
2760 ..CCGCACACATTTTACAGACTTCTAGTGTCTCCATGGGAATCCCTCA 2807
842 SerGlnSerValGlnSerIleArgProPro.TyrAsnArgAlaValSerI 858
2808 ACTCAACTCATTTCTGTCTGTGATTTCTCTACAAACAGACACTTCCC 2857
858 eu.....AspSerProValSerValGlySerSerProPro 869
2858 TCAGTCACATCACCAGCAGACATCATGCTCAGCAACAGCAGCAACTCAGCC 2907
870 ValLysAsnIleSerAlaPheProMetLeuProLysGlnProMetLeuG1 886
2908 GGCACAGAGACTGACGCTTCCCGACCTTCCCAAG..... 2943
886 yglYAsnProArgMetMetAspSerGlnGluAsnTyrGlySerSerMetS 903
2944 .....TTCACACACATAGCACACAGCTGCTCTCTCTT. 2976
903 LysGlyProAsnArgAsnValThrValThrGlnThrProSerSerGlyAsp 919
2976 ..... 2976
920 TyrGlyLeuProAsnSerLysAlaGlyArgMetGluProMetAsnSerAs 936
2977 .....GACATCAAGGAGGAGGAGGATGGCCCATTAAGAT.. 3012

```

```

936 nSerMetGlyArgProGlyLysPtyrAsnThrSerLeuProArgProa 953
3013 .....TACTCAGATGACTGAGGAAA..... 3033
953 lalaGlyLysSerIleProThrLeu.....ProLeuArgSerAsn 966
3034 .....GGAGGGAAGTCCAGCAGTTCATGACATGCAGATATGATGTT 3078
967 SerIleProGly.....AlaArgProValLeuGlnGlnGlnGlnG1 980
3079 CTAGTTCCTGGAATTAAGTGGCAGAGAAAATCCTGCTAGTCTACAGAT 3128
980 nMetLeuGlnMetArgProGlyLysIleLeuProMetGlyMetGlyAlaAsp 997
3129 GTACATTAATACCAACCAACA.....GGAGGTGATC 3160
997 roTyrGlnAlaAlaAlaSerAsnGlnLeuGlySerTyrPro..... 1011
3161 ATAGGGGACAGCCAGTTCGACAGTGTTTAGTGTGCTGATATTTT 3210
1012 AspGlyMetLeuSerMetGlnGlnValSerHisGlyThrGlnAsnArgPr 1028
3211 GATGGAAGAAAGATATATGCCCCAAATA..... 3237
1028 oleuLeuArgAsnSerLeuAspAspLeuValGlyProProSerAsnLeuG 1045
3238 .....TTAGAAAGCTCAGCT.....ATGAATAGACTCCAGGGAATCAGA 3277
1045 LnuGlyLysSerAsp 1049
3278 AAGCAGCATATGAT 3291

```

seq_name: /cgn2_6/ptodata/1/lna/6B.COMB.seq:US-09-107-847-1

seq_documentation_block:

; Sequence 1, Application US/09107847

; Patent No. 6100062

; GENERAL INFORMATION:

; APPLICANT: DUCKWORTH, DAVID

; APPLICANT: MICHALOVICH, DAVID

; TITLE OF INVENTION: NOVEL USE

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESS: Ratner & Prestia

; STREET: P.O. Box 980

; CITY: Valley Forge

; STATE: PA

; COUNTRY: USA

; ZIP: 19482

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107.847

; FILING DATE: 30-JUN-1998

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: EP 97304996.8

; FILING DATE: 08-JUL-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Prestia, Paul F

; REGISTRATION NUMBER: 23,031

; REFERENCE/DOCKET NUMBER: GH-30003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-407-0700

; TELEFAX: 610-407-0701

; TELEX:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5715 base pairs

; TYPE: nucleic acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-107-847-1

alignment_scores:
  Quality: 313.00      Length: 1172
  Ratio: 0.607        Gaps: 54
  Percent Similarity: 44.027  Percent Identity: 19.795

alignment_block:
US-09-041-994-2 x US-09-107-847-1 ..

Align seg 1/1 to: US-09-107-847-1 from: 1 to: 5715

34  GluIysArgArgGluGluGlnIuSerIysTyrIleGluGluLeuAlaG1 50
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
378  GAAAGAAACGTAGCATCAATTTAATGTCATTAAGAAGACGTGGATC 427
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
50  uLeuIleSerAlaAsnLeuSerAspIleAspAsnIleAsnValLysProa 67
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
428  CATGCTCTCTGTATGCTAGA.....AAGATGG 456

67  sPlvsCyAlaIleLeuLysGluThrValArgIleIleArgIleLys 83
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
457  ACAAACTTACTGTCTGCAGAAAGCATGATTTTTCGAAACATGAAA 506
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
84  GluGlnGlyThrIleSerAsnAspAspValGlnLysAlaAspVal 100
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
507  GAA.....ATCAGTCGACAGTCAGATCTGTGAAATTCGACAGACTG 550
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
100  lSerSerThrGlnGlnGlyValIleAspLysAspSerLeuGlyProLeuL 117
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
551  GAAACCTTACA.....TTCCTTAGTAAGAGAGTTTACACATTTAA 591

117  euLeuGluAlaLeuAspGlyPheLeuPheValAlaAsnArgGluLysAsn 133
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
592  TGTTAGAGCTCTGTGATGTTTTTTTAGCATCATGACAGATGGAAAGC 641

134  lIleValPheValSerGluAsnValThrGlnIleuGlnIleuGlnGlnG1 150
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
642  ATAATATATGTGTGAGAGTGTACTTCACTTAGAACATTTCCATC 691

150  uAspLeuValAsnThrSerValTyrAsnIleLeuHiscLysGluAspArgL 167
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
692  TGATCTTGTGATCAAGATATATTATCCAGAGAGGGAACATTT 741

167  yAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrp 183
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
742  CAGAGCTTATTAATCTCTCTACTCATCTGCTGGAAGTGAATTCATTA 791

184  ThrAsnGluProGlnArgGlnLysSerHis...ThrPheAsnCyArgMe 199
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
792  ACCCAGAAATTTTAATCAAAAAATCAGTTAGAAATCTGTGTGCACAT 841

199  tLeuMetLysThr.....ProHisAsp..... 206
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
842  GCTGCGAGAACAAATAGACCAAGAGCCATCTACTATGAATGTGTA 891

207  .....IleuGluAspIleAsnAlaSerProLys 216
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
892  AATTATATGAATTTCAATCTTTAAACAGTATCTCTTCACGACATC 941

217  .....MetArgGlnArgTyrGluTh 223
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
942  AATGTTTTGAAGAACTATACACGACACATAGCCATCTTATGAAAGA 991

223  tMetGlnCysPheAla.....LeuSerGlnProArgAlaMet 236
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
992  TAGAGTTTGTGTTGTAGCTACTGTGACGTAGCTACACCTCAGTTCATCA 1041

236  etGluGlnGlyGluAspLeuGlnSerCysMetIleCysValAlaIleArg 252

```

```

1042  AGGAA.....ATGTGC..... 1052
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
253  lIleThrGlnGlyLysArgThrPheProSerAsnProGluSerPheIleTh 269
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1053  .....ACTGTTGAGAACACCAATGAAAGATTTACATC 1084

269  rArgHisAspLeuSerGlyLysValValAsnIleAspThrAsnSer... 284
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1085  TAGACATAGTTTAGAATGGAAGTTTCTGTTCTAGATCAGAGCGACAC 1134

285  .....LeuArgSerSerMetArgPro 291
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1135  CCATATAGGCTATTTGCCATTGTAAGTTCTGGAAACATCA..... 1175

292  GlyPheGluAspIleIleArgArgCysIleGlnArgPhePheSerLeuAs 308
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1176  GGCTATGAT.....TCTATTCATGTGCA 1198

308  nAspGlyLysSerTrpSerGlnLysArgHisTyrGlnIleuAlaTyrLeuA 325
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1199  TGACCTAGAAATTTGGCAAAATGTCATGACACCTTAATGCCAATAT... 1244

325  snGlyHisAlaGlnThrProValTyrArgPheSerLeuAlaAspGlyThr 341
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1245  ..GGGAAAGCCAAATCATGTTATTTAGGTTCTGACTAAGGGGCACAG 1292

342  lIleValThrAlaGlnThrLysSerLysLeuPheArgAsnProValThrAs 358
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1293  TGGATTTGGCTTACAGCT..... 1310

358  nAspArgHisGlyPheValSerThrHisPheLeuGlnArgGluGlnAsnG 375
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1311  .....CATTTATATATCACTTACCAT.....CAGTGAATTT 1341

375  lYTyArgProAsnProAsnProValGlyGlnGlyIleArgProMet 391
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1342  CA.....AGGCCAGAGTTT 1355

392  AlagIlyAsnSerSerValGlyGlyMetSerMetSerProAsnGlnG1 408
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1356  ATTGTTTGTACTACACATGTAAGTATATGCA.....GA 1390

408  yLeuGlnMetProSerSerArgAlaTyrGlyLeuAlaAsp.....ProS 423
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1391  AGTTAGGCGCTGAAGACGACGAGAACTTGGCATTGAAGAGCTCTTCTCG 1440

423  erThrThrGlyGlnMetSerGlyAlaArgTyrGlyGlySerSer..... 437
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1441  AGACAGCTGTGACAAAGC.....CAAGATTCGTGGGTCAGATTAATCT 1484

438  ..AsnIleAlaSerLeu.....ThrPr 444
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1485  ATAAACACAGTCACATCAAGAAAGCATTTGAAAGTTTGCATCACAGCC 1534

444  oGlyProGlyMetGlnSerProSerSerTyrGlnAsnAsnAsnTyrGlyL 461
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1535  AACCCTTGTGCTCTCTCGAGTTCAAGAAATCATCTCACACGCGCG 1584

461  euAsnMetSerSerProProHisGlySerProGlyLeuAlaProAsn... 476
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1585  TC.....TCAGACCTTCTCTACACCAACCAAGATCCGACGAT 1625

477  .....G1 477
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1626  ACGACACTCCACCCAGCAGCATTTACCAGCTCATGGAAGATGTGTCA 1675

477  nGlnAsnIleMetLysSerProArgAsnArgGlySerProLysIleAlas 494
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1676  AAGAGGTCATCATTTAGTACTAGTCCATTAATTCACAGTCTGTGGTT 1725

494  erHisGlnPheSerProValAlaGlyValHisSerProMetAlaSerSer 510
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```


[illegible]

743	AspArgAspAspProSerAspAlaLeuSerTyrSGluLeuGlnProGlnI	759
2493TCACGACATTGTCAGTACGACGACACACCA	2527
759	IGluGlyValAspAsnLysMetSerClnGlyThrSerSerThrIleProS	776
	:::	
2528	GCAG.....ACGTCCAGACGACGACGCTCACTTCAGTTCAGCAACAT	2571
776	erSerSerGlnGluLysAspProLysIleLysThrGluThrSerGluGlu	792
2572	CTCAGGCTCAGCTGACCCAGCA.....	2594
793	GlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLeuThrSe	809
2594	2594
809	rSerAspPheTyrAsnAsnSerIleSerSerAsnGlySerHisLeuGlyT	826
2594	2594
826	hTrpGlnGlnValPheGlnGlyThrAsnSerLeu..GlyLeuLysSer	841
2595	..CCGAAACAAATTTTACAGACTTCAAGTTCCTCAATGGAAATCCCA	2642
842	SerGlnSerValGlnSerIleArgProPro..TyrAsnArgAlaValSerL	858
	:::	
2643	ACTCACACTCATTTCTCTCTGCTGCTGCTCTCTCACACAGACGACCTTCC	2692
858	eu.....AspSerProValSerValGlySerSerProPro	869
2693	TTCAGTCACATCACACGACCAATCATGCTCTCCGCAACGACGACCACTCACCC	2742
870	ValLysAsnIleSerLysIlePheProMetLeuProLysGlnProMetLeuG	886
2743	GGCACACGACTGACACTTCCCGGACCTTCCAAAG.....	2778
886	gLyAsnProArgMetLeuMetLeuSerGlnGluAsnTyrGlySerSerMetG	903
2779TTCACACAGTACGACACGCTGCTCTCTT..	2811
903	LysGlyProAsnArgAsnValThrValThrGlnThrProSerSerGlyAsp	919
2811	2811
920	TrpGlyLeuProAsnSerLysAlaGlyArgMetGluProMetAsnSerAs	936
	:::	
2812GACATCAAGGAGGAAGGATGGCCCATTTAAGAT..	2847
936	nSerMetGlyArgProGlyGlyAspTyrAsnThrSerLeuProArgProA	953
2848TACTCAGATGACCTGAGGAAA.....	2868
953	IleLeuGlyGlySerIleProThrLeu.....ProLeuArgSerAsn	966
2869GGAGGAAAGTTTCCAGCAGTTTATCATGAGATCAGATGACGTGTT	2913
967	SerIleProGly.....AlaArgProValLeuGlnGlnGlnGlnGlnG	980
	:::	
2914	CTAGTTCTCGGAATTAGTGGCAGAGAAATGCTCCATGAGCTTACAGAT	2963
980	nMetLeuGlnMetArgProGlyGlyIleProMetGlyMetGlyAlaAsnP	997
2964	GTCATTTAAATACCCACCCAGCA.....GGAGCTGATC	2995
997	roTyrGlyGlnIleAlaIleAsnAsnGlnLeuGlySerTyrPro.....	1011
2996	ATAGGGGCACACGCCATTTGACAGCTTTTACGCTGATATTTT	3045
1012	AspGlyMetLeuSerMetGlnGlnValSerHisGlyThrGlnAsnArgPr	1028
3046	GATGGAAAAAGATATATTGGCCAAATA.....	3072

```
1028 cLeuLeuArgAsnSerLeuAspAspLeuValGlyProProSerAsnLeuG 1045
      |||||.....:
3073 ...TTAAGAAGCTCAGCT.....ATGAATGACCTCCAGGAAATCAGA 3112
      :|||:
1045 luGlyGlnSerAsp 1049
      :|||:
3113 AAGGCACTAATCAT 3126

seq_name: /cgn2.6/ptodata/1/lna/5b_COMB.seq:US-08-816-693A-1

seq_documentation_block:
; Sequence 1, Application US/0881693A
; Patent No. 5874241
;
; GENERAL INFORMATION:
; APPLICANT: Takahashi, Joseph S
; APPLICANT: Turek, Fred W
; APPLICANT: Plinto, Lawrence H
; TITLE OF INVENTION: Clock Gene and Gene Product
; NUMBER OF SEQUENCES: 53
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz
; STREET: Two Prudential Plaza, Suite 4700
; City: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/816.693A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5874241thrup, Thomas E
; REGISTRATION NUMBER: 33,268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5400
; TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7498 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 389..2954
;
; US-08-816-693A-1

alignment_scores:
      Quality: 309.50      Length: 1138
      Ratio: 0.614      Gaps: 51
      Percent Similarity: 44.288      Percent Identity: 19.947

alignment_block:
US-09-041-994-2 x US-08-816-693A-1 ..

Align seg 1/1 to: US-08-816-693A-1 from: 1 to: 7498

34 GluLysArgArgArgGlnGlnGlnSerLysTyrIleGlnGlnLeuAlaG1 50
      |||||.....:
515 GAAAGAAGACGTAGAGATCAATTCAATGCCATTAAGAGAGAGCTGGGTC 564
      :|||:
50 uLeuLeuSerAlaAsnLeuSerAspIleAspAsnPhenAlaLysPro 67
      :|||:
565 TATGCTTCTCTGTTACGCGAGA.....AAGATGG 593
      :|||:
67 splysCyAlaIleLeuLysGluThrValArgGlnIleArgGlnIleLys 83
```

```
594 ACNAGTCTACTGTTCTACAGAAAGACATGATTTTTCGCAACATPAA 643
      |||||.....:
84 GluGlnGlyLysThrIleSerAsnAspAspValGlnLysAlaAspVa 100
      |||||.....:
644 GAG.....ACCACGTGCACAGTCAGATCTGTAGTATCCACAGAGCTG 687
      :|||:
100 lSerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeu 117
      :|||:
688 GAAACCCACA.....TTCCTTAGTAATGAGAGATTACACAGTTAA 728
      :|||:
117 euLeuGlnAlaLeuAspGlyPheLeuPheValAlaAsnArgLysGlnG1 133
      :|||:
729 TGTTAGAGGCTCTGTGATGTTTTTTTAAAGCATCATCAGACATGAGAGT 778
      :|||:
134 lIleValPheValSerGlnAsnValThrGlnIleGlnLysTyrLysGlnG1 150
      :|||:
779 ATAAATATATGATCTGACAGAGTAACTGCTGACTTGAACATTTACCATC 828
      :|||:
150 uAspLeuValAsnThrSerValTyrAsnIleLeuHISGlnGlnAspArgL 167
      :|||:
829 TGATCTGTGATCAAACTATATTATTTATCCACAGAGGAGACAATY 878
      :|||:
167 yAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrp 183
      :|||:
879 CAGAGGTTTATPAGATCTCTACTCATCTGCTGAAAGTACATTA 928
      :|||:
184 ThrAsnGluProGlnArgGlnLysSerHis...ThrPheAsnArgme 199
      :|||:
929 ACCCTGAGTACTTAAATCAAAAATCAGTAGATTTCTGTGACAT 978
      :|||:
199 lLeuMetLysThr.....ProHISp..... 206
      :|||:
979 GCTTCGAGGAAACAATAGACCAAGACCATCCACTATGAAATATGTA 1028
      :|||:
207 .....lIleLeuGluAspIleAsnAlaSerProGlu 216
      :|||:
1029 GATTTATAGAAATTTTAAATCTTTAACCGAGTATCACTTCAACAC 1078
      :|||:
217 .....MetArgGlnArgTyrGluThr 223
      :|||:
1079 AATGTTTGAAGAACTATACAAACGACACATAGGCTCTTATGAAAGA 1128
      :|||:
223 lMetClnLysPheAla.....LeuSerGlnProArgAlaMetm 236
      :|||:
1129 TAGACTGTGTTTGTAGCTACTGATTCAGATTAGCTACACTAGTTCA 1178
      :|||:
236 etGluGlnGlyGluAspLeuGlnSerCysMetIleCysValAlaArgArg 252
      :|||:
1179 AGGA.....ATCTGT..... 1189
      :|||:
253 lIleThrThrGlyGluArgThrPheProSerAsnProGlnSerPheIleTh 269
      :|||:
1190 .....ACTGTGAGAACCAACCAATGAAAGATTATACATC 1221
      :|||:
269 lArgHisAspLeuSerGlyLysValValAsnIleAspThrAsnSer... 284
      :|||:
1222 TAGACACAGTTTAGAATGAAAGTTTCTATTATTAGATACAGGACACAC 1271
      :|||:
285 .....LeuArgSerSerMetArgPro 291
      :|||:
1272 CAATATAGGCTATTGGCATTTGAGTCTTGGGAACATCA..... 1312
      :|||:
292 GlyPheGluAspIleIleArgArgCysIleGlnArgPhePheSerLeuAs 308
      :|||:
1313 GGCTATGAT.....TACTATCATGTGGA 1335
      :|||:
308 nAspGlyGlnSerTrpSerGlnLysArgHisTyrGlnGlnAlaTyrLeuA 325
      :|||:
1336 TGACCTAGAAAATCTGCCAAAATGTCACGAGCACTTATATGCAATAT... 1381
      :|||:
325 snGlyHisAlaGluThrProValTyrArgPheSerLeuAlaAspGlyThr 341
      :|||:
```

```

1382 ..GGAAAGGCAATCGTGTACTATGATTCTGACCAAGGCCACAG 1429
342 ILeValThrIaGlnThrLysSerLysLeuPheArgAsnProValIThrAs 358
      ::::|
1430 TGGATATGGCTTCAGACT..... 1447
358 nAspArgnHISGlyPheValSerThrHISpHeuGlnArgGluGlnAsnG 375
      ||| ::::|
1448 .....CATTTATATTTACTTACCAT.....CAGGGCAATT 1478
375 LyTyArgProAsnProAsnProValGlyGlnGlyIleArgProPmet 391
      ::::|
1479 CA.....AGCCGAGACTTC 1492
392 AlAGLyAsnSerSerValGlyGlyMetSerMetSerProAsnGlnG 408
      ||| ::::|
1493 ATGTGTTGTACTCAGACTGTAGTAAGTTATGA.....GA 1527
408 YLeuGlnMetProSerSerArgAlaTyrgLYLeuAlaAsp.....ProS 423
1528 AGTTAGGGCTGAAAGACGGCGAGACACTGGCATTTGAAGACTCTCTCTG 1577
423 eThrThrGlnGlyGlnMetSerGlyAlaArgTyrgLYSer..... 437
      ::::|
1578 AGACAGCTGCTGACAAAGC.....CAAGATTCTGGCTGACAAATCGT 1621
438 ...AsnIleAlaSerLeu.....ThrPr 444
      ||| ::::|
1622 ATCAACACAGTGTGATCTCAAGAGACACTGGAAAGTTTGATCACACCCC 1671
444 OGlyProGlyMetGlnSerProSerSerTyrgIAsnAsnAsnTyrgLYL 461
      ||| ::::|
1672 AACTCTCTTGCCCTCTAGAAAGCTCAGAAAGTCATCTCACACCCGCG 1721
461 euAsnMetSerSerProProHISGlySerProGlyLeuAlaProAsnGln 477
      ::::|
1722 TC.....TCAGACCTTCTCTCCACACCGCAAGATCCCTACTGAT 1762
478 GlnAsnIleMetIleSerProArgAsnArgGlySerProLysIleAla 494
      ::::|
1763 ACTAGC.....ACTCTCCACAGACAG..... 1783
494 rHISGlnPheSerProValAlaGlyValHISserProMetAlaSerSerg 511
1784 .....CATTTGCCAGCTCATGAAAGA 1805
511 LysAsnThrGlyAsnHISserPheSerSerSerSerLeuSerAlaLeuGln 527
      ::::|
1806 TGACACAGCGGAGGTGCTCTTCAGCAGTCAGTCCATTAAC..... 1846
528 AlAlaIleSerGluGlyValGlyThrSerLeuLeuSerThrLeuSerSer 544
      ||| ::::|
1847 .....TCCAGTCAGTGTGCTCCATCATTTAACAG..... 1876
544 OGlyProLYLeuAspAsnSerProAsnMetAlaIleThrGlnProSerL 561
1877 .....CCAGCATGTCTCAAGCTGCATAATTACCAATT.....CCACAG 1916
561 yValSerAsnGlnAspSerSerLysSerProLeuGlyPheTyrgYAspGln 577
      ::::|
1917 GCATGTTCACAGTTTCACTTTTTCAGCTCAGTTAGA..... 1951
578 AsnProValIleuSerSerMetCysGlnSerAsnSerArgAspHISLeu 594
1952 .....GCCATGCAGCATCTTAA 1968
594 rAspLYGlnSerLysGlnSerSerValGlyAlaGlnAsnGlnArg 611
1969 AGACACAGCTAGAGCAGCGACACGATGATAGAGCAATAATTTCATGGGC 2018
611 LyProLeuGlnSerLysGlyHISLysLysLeuLeuGlnLeuLeuThrCys 627
2019 AGCAAGAAAGAACTAAGAAATTCAGAGCAACTTCAGATG..... 2059

```

```

628 SerSerAspAspArgGlyHISSerSerLeuThrAsnSerProLeuAsp 644
2059 ..... 2059
644 rSerCysLYGlnSerSerValSerValIThrSerProSerGlyValSers 661
2059 ..... 2059
661 eSerThrSerGlyGlyValSerSerThrSerAsnMetHISGlySerLeu 677
2060 .....GTCATGTCAGCAGCG 2074
678 LeuGlnGlnLysHISArgIleLeuHISLysLeuLeuGlnAsnGlyAsn 694
2075 CTACAG.....ATGTTTTCAGCAATCAAC 2101
694 rProAlaGlnValAlaLysIleThrAlaGlnAlaThrGlyLysAspThr 711
2102 .....CCTGCAATTGATTTTG 2117
711 eSerIle...ThrSerCysGlyAspGlyAsnValValLysGlnGln 726
2118 GTTCTGTCAACTTCTCTCGAAATTCATATATC.....CACAG 2158
727 LeuSerProLYSlySlyGlnAsnAsnAlaLeuLeuArgTyrgLeuLeu 743
2159 CTCACACT..... 2167
743 pArgAspAspProSerAspAlaLeuSerLYGlnLeuGlnProGlnValG 760
2168 .....GTAATATCAAGCCAGGTT. 2188
760 LngLYValAspAsnLysMetSerGlnCysThrSerSerThrIleProSer 776
2189 .....GTGCCCT 2197
777 SerSerGlnGlnLysAspProLYSlyIleLysThrGlnThrSerGlnGln 793
2198 AACCCAGGTTTCAG..... 2209
793 ySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLeuThrSers 810
2210 ..AGTGCATATC..... 2221
810 eAspPheTyrgAsnAsnSerIleSerSerAsnGlySerHISLeuGlyThr 826
2222 .....AGCACAGCCACACATG..ATA 2242
827 LysGlnGlnValPheGlnGlyThrAsnSerLeuGlyLeuLysSerSerg 843
2243 CAGCAACAGACTTTTACAAAGTATACATCA.....ACTCGCA 2277
843 nSerValGlnSerIleArgProProTYrgAsnArgAlaValSerLeuAsp 860
2278 GAGTCAACAAGATGTATAGATGACGACACACTCAGACAGCTCTTCCAA 2327
2328 GTCAGACACGAGACTCTCAGACCCCACTGTACAT..... 2365
877 ProMetLeuProLYSglnProMetLeuGlyGlyAsnProArgMetMetAs 893
2366 ACGATGTGATTTTCCAGCTGCAGCTGG..... 2395
893 pSerGlnGlnAsnTyrgLYSerSerMetGlyGlyProAsnArgAsnValT 910
2396 ..AGCATGTCCAGATTCATCCAGTATG...CCACAGAAACAGTACCAGA 2441
910 hValThrGlnThrProSerSerGlyAspTyrgLYLeuProAsnSerLys 926
2442 GTGCTACAGTCACTACGTTCACTCAGACAGACAGATTAAGATTTTTCAA 2491

```

```

927  AAGlAAGTmetLurPmetAanSerAanSermetGlyrrProGlygl 943
2492  GGT.....CAGCACTTGTGGACCAATTAGTACTGCTCTGTAGC 2532
943  yAspTyraSThSerLeuProArgrProAlaLeuGlySerLe... 958
2533  TTGTGGGGCGGTATGTCATGACAGTACCATCATTTATGTGGTCAGGTGTGA 2582
959  ....ProThLeuProLeuAArgSerAanSerLeIleProGlyAlaArg 972
2583  CTGGCTATCTTACTTCTGCCACACACACAGCAGGCGACAG...ACATTGA 2629
973  ProValLeuGlnGlnGlnGlnGlnMetLeuGlnMetArProGlyGlu1 989
2630  TCGGTAAACACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 2679
989  eProMetGlyMetGlyAlaAan.....ProTyGlyGlnA 1001
2680  GCAACAACAACACACAGAGTTCACAGAAACAGCAGCTTCCTCAGTTGAC 2729
1001  IaLaIaIaSerAanGlnLeuGlySerTrpProAspGlyMetLeuSerMet 1017
2730  AGCCACTGTACGGCCACAGCTGGCCAGCCACACAGCAGTCTTACACACA 2779
1018  GlnGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAanSer.. 1033
2780  TCTAGTGTGCTCCACGGGAATCCTTCGCACACAGCTCATCTCTGCTGC 2829
1034  ....LeuAspAspLeuValGlyProProSerAan..... 1043
2830  CTTCCACTATCAACAGACAGCATCTTCCTCTGACACACCAGACAAACC 2879
1044  .....LeuGlnGlyGlnSer 1048
2880  AGCCTCAGCAGCAACAGCAGCTCTCTGCGACAGAGACTGCAGCTGACT 2929
1049  AArgGlnAArgAlaLeuLeuAArgGlnLeuHisThr...LeuLeuSerAan.T 1064
2930  GACCTTCACAGGTCCAGCCACAGTACGACACACACTTCTCTGTGCAT 2979
1064  hrAspAlaThrGlyLeuGlnGlnGlnLeuAspArgAlaLeuGlyIleProGln 1080
2980  GCGAGAGGAAGGGAGTGGCCAGAAAGAAATGCTGATGGCATGCGGTCA 3029
1081  LeuValAanGln 1084
3030  GAAGTTGAACAG 3041

seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:US-08-885-291-1

seq_documentation_block:
Sequence 1, Application US/08885291A
Patent No. 6057125
GENERAL INFORMATION:
APPLICANT: Takahashi, Joseph S.
APPLICANT: Turek, Fred W.
APPLICANT: Pinto, Lawrence H.
TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
FILE REFERENCE: 0290-5
CURRENT APPLICATION NUMBER: US/08/885_291A
CURRENT FILING DATE: 1997-06-30
EARLIER APPLICATION NUMBER: 08/816_693
EARLIER FILING DATE: 1997-03-13
NUMBER OF SEQ ID NOS: 55
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 7498
TYPE: DNA
ORGANISM: Mus musculus
US-08-885-291-1
alignment_scores:

```

[illegible]

1121ProMetGInGly 1122
7517 GCCCCAGGGGACCCAGGGCTACAGCCACCACTACCTGCGAGGTCAGG 7566
1125 1yPheHisLeuInGInGInSerProSerPheAsnSerMet..... 1138
7567 GGGTCCACTTCCAATCCAGCCATGCAGCAATGATATCCCAATGCAGGGCGG 7616
1139MetAsnGInMetAsnGInGInGlyAsnPh 1148
7617 GTTCAGAGGGCTGGCGTCGCCAGCAGCAACCAACAGCAGCAACTCCAGCC 7666
1148 eProLeuGInGlyMetHisProAlaAlaAsnIleMetArgProArgThr 1155
7667 ACCCAATGGAGGATGAGCCCCAGCGCTCAGCAGTGAATGAATGAACCA 7716
1165 snThr...ProGlyGInLeuArgMetGInLeuGInGInArgLeuGInGly 1180
7717 ACACCACTGCCCTTCACAAATTCGCA.....GACATCTTGAGACA 7754
1181 GInGInPheLeuAsnGInSerArgGInAlaLeuGInLeuGInMetGInAs 1197
7755 CACCAATGATGACACAGCAGCAGCA.....CAGGAGCAGG 7792
1197 nProThrAlaGlyGlyAlaAlaValMetArgProMetMetGInProGIn 1214
7793 GCCAGAAATAGGCGCTGGAAATGGCCACCACTAACCGATTCCAGCAACCC 7842
1214 InGlyPheLeuAsnAlaGInMetValAlaGInArgSerArgGInLeuLeu 1230
7843 AAGGAATTGGCTAACCCACCACAGCGCCAGCAGCG.....ATG 7880
1231 SerHisHisPheArgGInGInArgAlaAlaMetMetMetGInGInGIn 1247
7881 CAGCATCACATGCACACAGTGCACACAGGAATATGGCAGCATAGGCCA 7930
1247 nGInGInGIn.....GInGInGIn 1254
7931 GCTTCCGCCAGCGCTTGAGGACAGAGCAGAGTCCAGTACAGCGCTATC 7980
1254 InGInGInGInGInGInGInGInGInGInGInGInGInGInGInAla 1270
7981 AGCAGGAGCTCCTTCAGCACACAGATGGGCTCCCTGTTCAGGCCAACCC 8030
1271 PheSerProPro.....ProAsnValThrAlaSerProSerMet 1283
8031 ATGAGCCCCCAGAGAGATATGCTCCCAATATCAGGCCAGTCCCACT 8080
1283 tAspGlyLeuLeuAlaGlyProThrMetProGInAlaProGInGInP 1300
8081 ACAAGGC.....CAGCAGA 8094
1300 nProGlyGIn...ProAsnTyrGInMetGInGInGInProAspProAla 1315
8095 TCCCTATATCTCTCCCAATCAAGTGCCTGCCACACCTGTCTCTCT 8144
1316 PheGlyArgValSerSerProProAsnAlaMetSerArgMetGIn 1332
8145 CCAGGCCACACTCCAGGCCGCCACATCTCAATCTTCCCAAGATGCA 8194
1332 yProSerGInAsnProMetMetGInHisProGInAlaAlaSerIleTyr 1348
8195 GCCTCAGCTTCTCCACACAGCTTCCCAAGATTCCTCCCACTC 8244
1349GInSerSerGInMetGlyTyrProSerGly 1359
8245 CTGGAGCTGTACCTGCCACAGGCCAACCCCATGGAACAAGGCATTGTGCC 8294
1360 AsnIleuAlaArgAsnSerSerPheSerGInGInGInPheAlaHisGInG 1376
8295 AGCCCGGACCAAAATTCATCACTGTTTCTAGCTT.....GCTAG 8332
1376 yAsnProAlaValTyrSerMetValHisMetAsnGlySer 1390

```
seq_name: /cgr2_6/ptodata/1/lna/PCTUS_COMB.seq:PCT-US95-04682-1
8333 CAATCCA.....GGCATGGCAACCTCCATGTCGACAGC 8366
:||||| :|||||:|||||:|||||:|||||
seq_documentation_block:
: Sequence 1, Application PC/TUS9504682
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
: TITLE OF INVENTION: FACTOR P300 AND USES OF P300
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
: STREET: Ten Post Office Square
: CITY: Boston
: STATE: MA
: COUNTRY: US
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/04682
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/227,536
: FILING DATE: 14-April-1994
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Holliday C. Heine, Ph.D.
: REGISTRATION NUMBER: 34,346
: REFERENCE/DOCKET NUMBER: DECI-308Xq999
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 542-2290
: TELEFAX: (617) 451-0313
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 9046 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1200..8441
: PCT-US95-04682-1

alignment_scores:
  Quality: 274.00      Length: 699
  Ratio: 0.828        Gaps: 29
  Percent Similarity: 47.353  Percent Identity: 23.176

alignment_block:
US-09-041-994-2 x PCT-US95-04682-1 ..
Align seg 1/1 to: PCT-US95-04682-1 from: 1 to: 9046

770 ThrSerSerThrIleProSerSerSerGlnGluTyrAspProTyrIleLeu 786
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6562 TCTGTCGTCACATGCAAGCAGCTGCCAGGAGACAAATGCCGTCGCCG 6611
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
786 sThrcIu...ThrsrGluGluGlySerGlyAspLeuAspAsnLeuAspA 802
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6612 TTCTGCTTAACATCATCAGCAGAGCTCCGCGC..... 6642
802 lallleuGlyAspLeuThrSerSerAspPheTyrAsnAsnSerIleSer 818
```

```
6643 .....AGCAACAGCTGCACAGCACCGAC...TACAGCAGGCCCAATATGC 6681
819 SerAsnGlySerThrIleLeuGlyThrIleGlnGlnValIlePheGlnGlyThrAs 835
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6682 TTTCGAGGAGGATGGCCAGCATGCAGCGAGCTGTGTGGTGGGACGAA 6731
835 nSerLeuGlyLeuTyrSerSerGlnSerValGln.....s 847
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6732 CAGGGGCTTCCTCCCTCCCACTCTCCCACTCCACAGCAGACCACTGGGCA 6781
847 erIleArgProProTyrThrAsnArg.....AlaValSerLeuAspSer 860
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6782 ACAGCCAAACACACCCCGCAGACGCCCACTTCAGCTCAGCCTCAGCCTA 6831
861 ProValSerValGlySerSerProPro.ValIleAsnIleSerAlaPhe 877
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6832 CCCCTCCCAATAGCATGCCACCTACTTCCCGGAGCTCAAGCTGCTGCG 6881
877 rometLeuProTyrGlnProMetLeuGlyAsnProArgMetMetAsp 893
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6882 CCTGTGTCGA..... 6892
894 SerGlnGluAsnTyrGlySerSerMetGlyGlyProAsnAlaGlnValTh 910
6893 .....GGGTAAAGCAGCAGGCGCAGGTGACCCCTCC..... 6922
910 rValThrGlnThrProSerSerGlyAspTyrGlyLeuProAsnSerIleAs 927
6923 .....AACCCCTCCTCA...GACTGCTACACCAACCCCTTTCACAGG 6959
927 lAgIArgMetGluProMetAsnSerAsnSerMetGlyArgProGlyGly 943
6960 CCCCACCTACAGCAGTGAATGGCAATGCAGATTGCAGAGCAGCGCA 7009
944 AspTyrAsnThrSerLeuProArgProAlaLeuGlyGlySerIleProTh 960
7010 GAGCAGCGCCAGATGGCCAGCTGCAATTTTCAAGGCCAATGCACAC 7059
960 r.....LeuProLeuArgSerAsnSerIleProGlyAlaArgProVal 975
7060 ACCAATGCCCGCCGATGATCCATGCCATGCCCGCCATGGATGATGACCCACT 7109
975 euGlnGlnGlnGlnGlnMetLeu.GlnMetArgProGlyGlyIleProme 991
7110 CCCATGACAGCAGAGTCCCATGTCAGTGGCATTTGGAGCGAGATGGCAGC 7159
991 tGlyMetGlyAlaAsnProTyrGlyGlnAlaAlaIleSerAsnGlnLeu 1008
7160 AGGATGCGACAGCACGCCA..... 7178
1008 lYserTrpProAspGlyMetLeuSerMetGlnGlnValSerHisGlyThr 1024
7179 ..CCCTGGAGCCAGAGAGATGCTCAGCCCACTACAGCTACAGCTGCGG 7226
1025 GlnAsnArgProLeuLeuArgAsnSerLeuAspAspLeuValGlyProPr 1041
7227 ATGCCAAGCGCCCATGATGTACAGTGGCCCGCAGCAT.....GGTCAACC 7270
1041 oSerAsnLeu..... 1044
7271 TTGTGAACATGCTCCCAACACGAGATTGGCCAGGTAGTATCAGCCAC 7320
1045 ....GluGlyGlnSerAspGluArgAlaLeuLeuAspGlnLeuHisThr 1059
7321 TCAAACACGAGCAGCTGTGTCTAACAAGCCTTACAAACCTTTTGGGAGCT 7370
1060 LeuLeuSerAsnThrAspAlaThrGlyLeuGlnGluIleAspArgAlaLe 1076
7371 CTCAGGTCTCCAGCTCTCCCTGCGCAGCAGCAACGAGCTTAGATGCT 7420
1076 u...GlyIleProGlyLeuValAsnGlnGlyGlnAlaLeuGlnProTyrSg 1092
```



```

67 splysCysAlaIleLeuLysGIuThrValArgGIuIleArgGIuIleLys 83
   ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
703 ACAAAACACACCGTGGAGAAAGTCATCGATTCTTGCAGAAACACAAAT 752
84 GIuGIuGIuLysThrIleSerAsnAspAspValGIuLysAlaAspVal 100
   ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
753 GAA.....GTCTACGACCAACAGAAATC.....TGTGCAT 784
100 IseSerThrGIuGIu...GIuValIleAspLysAspSerLeuGIuProL 116
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
785 CCAGCAGGACGTGGAAGCCATCATCTCTCAGTAACGAAATTCACCCACC 834
116 euLeuLeuGIuAlaLeuAspGIuPheLeuPheValAlaAsnArgGIuAla 132
   ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
835 TCATGTTGGAGGCATTCAGTGCCTTCGTCATCGTCGTGACAAACAGACGCG 884
133 AsnIleValPheValSerGIuAsnValThrGIuIleLeuGIuIleTyrLysGIu 149
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
885 ACATCATCTTATGTGTGCGACAGTATCACACCTCTCTTGACATTTACC 934
149 nGIuAspLeuValAsnThrSerValTyrAsnIleLeuHisGIuGIuAspVal 166
   ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
935 GGCGGATGTCTAGATGATGACAGACTTGTAAATTTCTTCCAGACAGACAGAC 984
166 rGIuLysAspPheLeuLysAsnLeuProLysSerThrValAsnGIuValSer 182
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
985 ATTCGCAACTTATTAATAATCTTTCTTCCAAATGCTT.....GTGACG 1028
183 TrpThrAsnGIuProGIuAsnArgGIuLysSerHisThr.....PheAs 196
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1029 GATTCGCCCTCCCGACAAATTCCTAAATCTGCACACAGATTTACAGTTTAA 1078
196 nCysArgMetLeuMetLysThrProHisAspIleLeuGIuAspIleAsnVal 213
   ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1079 TTGCCATCTTCTCAGAGGC.....AGCT 1101
213 IseSerProLumetArgGIuArgGIuThrMetGIuIleLysPhe..... 227
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1102 TGAACCCAAAGAAATTTCCAACTTACGAATATTAATAATTTGTAGGAAT 1151
227 ..... 227
1152 TTTCGCTTACACAAATGTGCTACGCCCTCCTGAATGCTTTAGACA 1201
228 .AlaLeuSerGIuProArgAlaMetMetGIuGIuLysAspLeuGIuAsn 244
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1202 CACCCCTTCAAGACC...TGCATGTACCCCTAGGAAAGACGTC.... 1244
244 eTCysMetIleCysValAlaArgArgIleThr.....ThrGIuGIu 257
   ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1245 ..TGCTTCAATGCCACCGTGGCGCTGGCAACCCCGAGTTCTTAAAGAA 1292
258 ArgThrPheProSerAsnPro...GIuSerPheIleThrArgHisAspVal 273
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1293 ARGTGTAAGTCGACGACAACTTTAGAGAAATTCACCTCGACGATAGCTT 1342
273 uSerGIuLysValAlaAsnIleAspThrAsnSer..... 284
   ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1343 GGAATGGAATTTTATTTCTGTGATCAGAGCTCTCCCAATCATATGAGAT 1392
285 .....LeuArgSerSerMetArgProGIuPheGIu... 294
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1393 ACCTGCCCTTTGAAGTACTTGGACCTCA.....GGCTACCAACTAC 1433
295 .....AspIleIleArgArgCysIleGIuArgPhePhe 305
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1434 TACCAATATGATGACCTGAGACTCTCTGGCCAGGTGCCACACAGATCTGAT 1483
305 eSerLeuAsnAspGIuLysSer..... 312
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1484 GCAGTTTGGCAAGGAGAGTCTGTACCGGTTTACCGGTTTCTAACCAAGGCG 1533

```

```

313 .....TrpSerGIuLysArgHisTyrGIuGIuAlaTyr..... 323
   ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1534 ACAGATGATTTGGTGGCAA...ACCCACTACTACATCTACCAACCA 1580
324 LeuAsnArgHisAlaGIuThrProValTyrArgPheSerLeuAlaAspGI 340
   ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1581 TGGAACTCCAAAGCCTGAGTTCACTGATGCACACACATCTAGTGTAGTTA 1630
340 yThrIleValThrAlaGIuThrLysSerLysLeu...PheArgAsnProVal 356
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1631 CGCAGATGTTCGATGGAAAGAGACAGAGAGCTGCTCGAAGAACCCAC 1680
356 aLThrAsnAspArgHisGIuPheValSerThrHisPheLeuGIuArgGIu 372
   ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1681 CCACAGAGCCCATGCAC..... 1697
373 GluAsnArgLysArgProAsnProValGIuGIu.....GI 386
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1698 .....CCCTCTGCAGTGAAGAAAGAAAGACTACAG 1726
386 yLleArgProPheMetAlaGIuLysAsnSerSerValGIuGIuMetSerTr 403
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1727 CCTAGAGCCTCCA..... 1739
403 eLseProAsnGIuGIuLeuGIuMetProSerSerArgAlaTyrGIuLeu 419
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1740 ..CAGCCCTTATATGCATTGACATG..... 1763
420 AlaAspProSerThrThrGIuGIuMetSerGIuAlaArgTyrGIuLys 436
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1764 .....GGCCG 1768
436 rSerAsnIleAlaSerLeuThrProGIuProLumetGIuSerProSer 453
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1769 CTCAGGTCCTCCACG...AGCCCTCTGCATCAGCCTCTCAAGAGAT 1815
453 eTyrGIuAsnAsnAsnArgLysLeuAsnMetSerSerProHisArgLys 469
   ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1816 CCCACAAAGTCCTACACACAGGCATG.....TCAGAACCCACCTTC 1856
470 SerProGIuLeuAlaProAsnGIuAsnIleMetIleSerProArgAs 486
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1857 ACTCCAAACCAAGCTGATGCTGAGAAACACACACAGCTTTCCAAGAC 1906
486 nArgLysSerProLysIleAlaSerHisGIuPheSerProValAlaArgLys 503
   ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1907 GCCAACCTTACCC.....CAGAGATTACACAGGACGAGGCG 1941
503 aHisSerProMetAlaSerSerGIuAsnThrGIuAsnHisSerPheSer 519
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1942 TC.....AGCCAGGCGACCAATGCCAGCTGCTGTGATTCTCTCAGCG 1985
520 SerSerSerLeuSer.....AlaLeuGIuAlaIleSerGIuGIuVal 533
   ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1986 TCCTGGACCTCACAAGCAACTCTGCTGACAGAGCTCCCTCAG... 2030
533 IGIuThrSerLeuLeuSerThrLeuSerSerProGIuProLysLeuAspVal 550
   ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
2031 ....ACCGGCTTG.....CAGAGTCCACCTGCTCCAGTACGACACAGT 2067
550 snSerProAsnMetAsnIleThrGIuProSerLysValSerAsnGIuAsp 566
   ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
2068 TTTTCAGCACAATTCAGCATGTTCCAGACCATTTAA..... 2102
567 SerLysSerProLeuGIuPheTyrCysAspGIuAsnProValGIuLys 583
   ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
2103 .....GACCAACTGAGCAAGGACGAC 2125
583 rMetCysGIuSerAsnSerArgAspHisLeuSerAspLysGIuSerLys 600
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
2126 GATCTGTGACGACCAACATTCGG.....TGGCAGC 2154
600 IuSerSerValGIuGIuAlaGIuAsnGIuArgGIuProLeuGIuSerLys 616

```

```

2937 .....GGAACTGGGCTAATCA 2954
906 AsnArgAsnValThrValThrGlnThrProSerSerGlyAspTrpGlyLe 922
      ::::| | | | | | | |
2955 GCCACACAGCTCTACACCTCCGCCAACCCACGC ..... 2987
922 wPrasnSerIysAlaIyArgMetGlnProMetAsnSerSerMetG 939
      |||::: .....| | | | | | | |
2988 .CCCTCCCGCAGGGTCGCGCGGTCA .....GGTCTTAAGGTCTCC 3030
      ::::| | | | | | | |
939 IyArgProGlyIyAspTyAsnThrSerLeuProArgProAlaLeuGly 955
      ::::| | | | | | | |
3031 AGCAGCCATCAGGTAATAC .....CCTACACTGGA 3062
956 GlySerIleProThrIleuProLeuArgSerAsnSerIleProGly 970
      ::::| | | | | | | |
3063 .....GTCCAGACGCAACACGAGCTTAACCAATGAAACGGGGT 3101

seq_name: /cgn2_6/plodeta/1/lna/5B.COMB.seq:US-08-480-473B-1

seq_documentation_block:
Sequence 1, Application US/08480473B
Patent No. 5882914
GENERAL INFORMATION:
APPLICANT: Semenza, Gregg L.
TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,473B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/053001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3736 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-480-473B-1

alignment_scores:
Quality: 258.00 Length: 915
Ratio: 0.604 Gaps: 41
Percent Similarity: 46.667 Percent Identity: 19.781

alignment_block:
US-09-041-994-2 x US-08-480-473B-1 ..
Align seg 1/1 to: US-08-480-473B-1 from: 1 to: 3736

1 MetSerGlyLeuGlyIyAsnLeuAsp.....ProLeuAlaSerAspSe 15
|||:::| | | | | | | |
29 ATGGAGGGCGCGCGCGCGGAACGAGCAAGAAAGATTAAGTTCTGAACG 78

```

```

2937 .....GGAACTGGGCTAATCA 2954
906 AsnArgAsnValThrValThrGlnThrProSerSerGlyAspTrpGlyLe 922
      ::::| | | | | | | |
2955 GCCACACAGCTCTACACCTCCGCCAACCCACGC ..... 2987
922 wPrasnSerIysAlaIyArgMetGlnProMetAsnSerSerMetG 939
      |||::: .....| | | | | | | |
2988 .CCCTCCCGCAGGGTCGCGCGGTCA .....GGTCTTAAGGTCTCC 3030
      ::::| | | | | | | |
939 IyArgProGlyIyAspTyAsnThrSerLeuProArgProAlaLeuGly 955
      ::::| | | | | | | |
3031 AGCAGCCATCAGGTAATAC .....CCTACACTGGA 3062
956 GlySerIleProThrIleuProLeuArgSerAsnSerIleProGly 970
      ::::| | | | | | | |
3063 .....GTCCAGACGCAACACGAGCTTAACCAATGAAACGGGGT 3101

seq_name: /cgn2_6/plodeta/1/lna/5B.COMB.seq:US-08-480-473B-1

seq_documentation_block:
Sequence 1, Application US/08480473B
Patent No. 5882914
GENERAL INFORMATION:
APPLICANT: Semenza, Gregg L.
TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,473B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/053001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3736 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-480-473B-1

alignment_scores:
Quality: 258.00 Length: 915
Ratio: 0.604 Gaps: 41
Percent Similarity: 46.667 Percent Identity: 19.781

alignment_block:
US-09-041-994-2 x US-08-480-473B-1 ..
Align seg 1/1 to: US-08-480-473B-1 from: 1 to: 3736

1 MetSerGlyLeuGlyIyAsnLeuAsp.....ProLeuAlaSerAspSe 15
|||:::| | | | | | | |
29 ATGGAGGGCGCGCGCGCGGAACGAGCAAGAAAGATTAAGTTCTGAACG 78

```

```

15  rArgLysArgLysLeuProCysAspThrProGlnGlnGlyLeuThrCys 32
   |||||:||||
79  TCGAAAGAAAG.....TCTCGAG 98

32  eGcLgLnLsArGrArGrGlnGlnIserLysTyrIleGlnLeu 48
   :|||:|||||:|||||:|||||:|||||
99  AFGAGCCAGATCTGCGAGAGTAAGATGTGAAGTTTATATACGCTT 148

49  AlaGlnLeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnVal 65
   :|||:|||||:|||||:|||||:|||||
149  GCTCAT.....CAGTTGGCAGCTCCACATTAATGATGATTCGA 186

65  sProAspLysCysAlaIleLeuLysGlnThrValArg.....GlnIleA 80
   |||||:|||||:|||||:|||||:|||||
187  TCTGTATAGGCGCTGTGATGAGGCTTACCATCATTTGCGGTGCA 236

80  rGgLIleLsGlnGlnGlnLysThrIleSerAsnAspAspValGln 96
   |||||:|||||:|||||:|||||:|||||
237  GGAACCTCTGATGCTGTGT...GATTGGATATTGAAGATGACATG... 280

97  LysAlaAspValSerSerThrGlnGlnValIleAspLysAspSerLe 113
   |||||:|||||:|||||:|||||:|||||
281  .....AAGCACAGAT 291

113  uGlyProLeuLeuGlnAlaLeuAspGlyPheLeuPheValValAsnA 130
   :|||:|||||:|||||:|||||:|||||
292  GAATGCTTTTATTGAAAGCCTGTGATGTTTGTATGTTCTCACAG 341

130  rGgLIAlaAsnIleValPheValSerGluAsnValThrGlnTyrLeuGln 146
   :|||||:|||||:|||||:|||||:|||||
342  ATGATGCGACATGATTTACATTTGATTAATGTGACAAATACATCGGA 391

147  TyrLysGlnGlnAspLeuValAsnThrSerValTyrAsnIleLeuHisG 163
   |||||:|||||:|||||:|||||:|||||
392  TTAACCTGTTGAACCTAAGTGCACACAGCTGTTGATTTACTACATCC 441

163  uGlnAspArgLysAspPheLeuLysAsnLeuProLysSerThrValAsnG 180
   |||||:|||||:|||||:|||||:|||||
442  ATGTACCACTGAGAAATGAGAAATGCTTACACACAGA.....ATG 485

180  Ly...ValSerThrAsnGlnProGlnArgGlnLysSer..... 192
   :|||:|||||:|||||:|||||:|||||
486  GCCTGTGTGAAAGGTAAAGAAACAACACACAGCAAGCTTTTCTTC 535

193  .....HisThrPheAsnCys.... 197
   :|||:|||||:|||||:|||||:|||||
536  AGAATGAAGTGTACCTTAAGTACGAGAGAAAGAACTATGAACATAAAGTC 585

198  .....ArgMetLeuMetLysThrProHisAspIleLeuGlnAspI 211
   :|||:|||||:|||||:|||||:|||||
586  TGCACATGAAAGGTATTGCACTGCACAGGCAACATTCACGTATATGATA 635

211  LeAsnAlaSerProGlnMetArgGlnArgTyrGlnThrMetLysCysPhe 227
   |||||:|||||:|||||:|||||:|||||
636  CCAACAGTAACCAACT.....CAGTGTGGG 661

228  AlaLeuSerGlnProArgAlaMetMetGlnGlnLysAspLeuGlnSe 244
   :|||:|||||:|||||:|||||:|||||
662  TATAGAAACCACT.....ATGAC 681

244  rCysMetIleCysValAlaArgArgIleThrThrGlnGlnArgThrPhe 261
   :|||:|||||:|||||:|||||:|||||
682  CTGTTGGTGTGATTGTGAAACCACTT.....CCTCAC 716

261  rOserAsnProGln.....SerPheIleThrArgHis 271
   |||||:|||||:|||||:|||||:|||||
717  CATCAAAATTTGAATTCCTTAGATACAGACATTTCTCTCAGTGCACAC 766

272  AspLeuSerGlyLysValValAsnIleAspThrAsnSerLeuArgSerSe 288
   :|||:|||||:|||||:|||||:|||||
767  AGCCTGATATGAATTTTCTTATTGTGTATGA.....AGAATTAC 807

```

```

288  rMetArgProGlyPhe.....GluAspIleIleArgArgCysIleGlnA 303
   :|||:|||||:|||||:|||||:|||||
808  CGAATTGATGGATATGAGCCAGAGAACTTTAGCCGCTCAATTTATG 857

303  rGpPhePheSerLeuAsnAspGlyInsSerThrPserGlnLysArgHisTyr 319
   :|||:|||||:|||||:|||||:|||||
858  AATATTATCATGCTTTGGACCTGTGATCATCTGACCAAAATCATAT... 904

320  GlnGlnAlaTyrLeuAsnGlyHisAlaGlnThrProValTyrArgPheSe 336
   :|||:|||||:|||||:|||||:|||||
905  ...GATATGTTACTAAAGAGCAAGTCAACACAGACAGATCAGATGCT 951

336  rLeuAlaAspGlyThrIleValThrAlaGlnThrLysSerLysLeuPheA 353
   |||||:|||||:|||||:|||||:|||||
952  TGCCAAAGAGGTGATATGCTGGTTGAAACTCAACCACTGATCATAT 1001

353  rGAsnProValThrAsnAspArgHisGlyPheValSerThrHisPheLeu 369
   |||||:|||||:|||||:|||||:|||||
1002  ATBACACCAAGAAATTCACACAGCTGATGTATGTGTAAATTACGTT 1051

370  GlnArgGlnGlnAsnGlyTyrArgProAsnProAsnProValGlnGln 386
   |||||:|||||:|||||:|||||:|||||
1051  ..... 1051

386  rYleArgProPheMetAlaGlyCysAsnSerSerValGlyGlyMet.... 401
   |||||:|||||:|||||:|||||:|||||
1052  .....GTGAGGTATTTATTC 1067

402  .....SerMetSerPro 405

1068  AGCAGACTGATTTTCCTCCCTCAACAACAGATGTGCTTAAACG 1117

406  AsnGlnGlyLeuGlnMetCProSerSerArgAlaTyrGlyLeuAlaAspPr 422
   :|||:|||||:|||||:|||||:|||||
1118  GTTGATCTTCAGATATGAAATGACTCAGCATTTACACAAAGTTGAAATC 1167

422  oSerThrThrGlnMetSerGlyAlaArgTyrGlyLysSerAsnI 439
   :|||:|||||:|||||:|||||:|||||
1168  AGAAGATCAAGTACCTCTTGGACAAACTTAAGAGAACCTGATGCTT 1217

439  LeAlaSerLeuThrProGlyProGlyMetGlnSerPserSerTyrGln 455
   :|||:|||||:|||||:|||||:|||||
1218  TAACTTGTGCGCCCGCCAGCGCTGCA.....GACCATATCTTA 1255

456  AsnAsnAsnTyrGlyLeuAsnMetSerSer..... 465
   :|||:|||||:|||||:|||||:|||||
1256  TCTTTAGATTTTGGCAGCAACGACACAGAAACTGATGACCAACTTGA 1305

466  .....ProPheHisGlySerProGlyLeu...AlaProAsnGlnGln 478
   |||||:|||||:|||||:|||||:|||||
1306  GGAAGTACCATTTATATATGATGATATGCTCCCTCCACCAACGAAAT 1355

479  .....AsnIleMetIleSerProArgAsnArgGlySerProLys 491
   |||||:|||||:|||||:|||||:|||||
1356  TACAGATATATAATTTGGCAATGCTCCA..... 1384

492  IleAlaSerHisGlnPheSerProValAlaGlyValHisSerPheMetAl 508
   :|||:|||||:|||||:|||||:|||||
1385  .....TTACCCACCCCTGMAACGCCAACCCACTTCG 1416

508  aSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerLeuSera 525
   |||||:|||||:|||||:|||||:|||||
1417  AAGTAGTCTGAC..... 1429

525  lAlaGlnAlaIleSerGlnGlyValGlyThrSerLeuLeuSerThrLeu 541
   |||||:|||||:|||||:|||||:|||||
1430  ....CTGCACTCAATCAAGAAGTTGCATTAATTA..... 1462

542  SerSerProGlyProLysLeuAsnAsnSerProAsnMetAsnIleThrGln 558
   :|||:|||||:|||||:|||||:|||||
1463  ...GAACCAAAATCCAGAG.....TCACGTGAACCTTCTTTTACCAT 1500

558  nPro.....SerLysV 562

```

```

1501 GCCCGAGATTCAGATCAGACACACCTAGTCTCCAGTGAACGACCTAGAC 1550
      |||
562 a1SerAsnGlnAspSerIysSerPro.....LeuGlyPheTyrCysAsp 576
      |||:::  ::::|  |||
1551 AAGATTACCTGAGCCTATAGTCCAGAGAAATATTGTTTATATGAGAT 1600
      |||:::  ::::|  |||
577 GlnAsnProValGlnSerSerMetCysGlnSerAsnSerArgAspHisLe 593
      ::::|  |||:::  ::::|  |||
1601 ACTGATATGCTCAATGAAATTCAGATTGGAATGGTAAAGAACTTTTTC 1650
      ::::|  |||:::  ::::|  |||
593 uSerAspIysGlnSerIysGlnSerSerValGlnGlyAlaGlnGlnA 610
      ::::|  |||:::  ::::|  |||
1651 TGAAGACACAGAAAGCAAG..... 1669
610 rglGlyProLeuGlnSerIysGlnHisIysIysLeuLeuGlnLeuThr 626
      ::::|  |||:::  ::::|  |||
1670 ..AACCATTTTCTAGTCTAGACAGACAGATTAGACTTGCAGATGTTAGCT 1717
      |||
627 .....CysSerSerAspAsp.....ArgGlyHisSerSerLe 637
      |||
1718 CCTATATCCCAATGATGATGACTTCAGTTACGTTCTCTGATCAGTT 1767
      |||
637 uThrAsnSerProLeuAspSerSer..... 647
      |||
1768 G.....TCCACCTTAGAAAGCAGTTCCGCAAGCCCTCAAGACCGCAAGTC 1811
647 ysgGlnSerSerValSerValIhrSerProSerGlyValSerSerSerThr 663
      |||
1812 CTCAAGACACAGCTTACAGTATTCACAGACAGACCTCAATACAAAGCCTACT 1861
664 SerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeuGlnGln 680
      ::::|  |||:::  ::::|  |||
1862 GCTAATGCCACACCTACACCTGCTGATGATTAATAAAGAGTGCAC 1911
680 uIysHisArgIleLeuHis...LysLeuLeuGlnAsnGlnIysAsnSerPro 696
      |||
1912 AAAAGACCGCTATGGAAGACATTAATAATATTGATTCATCTCCTCA 1961
696 IagIuValAlaIysIleThrAlaGlnAlaIleThrGly.....LysAsp 709
      ::::|  |||:::  ::::|  |||
1962 CCCACATACATTAAGAACTACTAGTCCACACATCACCATATATAGAGAT 2011
710 ThrSerSerIleThrSer.....CysGlyAspGly..... 719
      |||
2012 ACTCAAAAGTCGACAGCCTCACCACAAACAGACGAAAGAGTCAATAGA 2061
719 ..... 719
2062 ACAGACAGAAAATCTCATCCAAAGACCCCTAACGCTTATCTGCGCTT 2111
720 .....AsnValValIysGlnGlnGlnLeuSerProIys..... 730
      ::::|  |||:::  ::::|  |||
2112 TGAGTCAAAAGAACTACAGTCTCTGAGGAAAGAACTTAATCCAAAGATCTA 2161
731 .....LysIysGlnAsnAsnAlaLeuLe 738
2162 GCTTTCAGCAATGCTCAGAGAAAGCGAAATGGAACATGATGATGTTCACT 2211
738 .....ArgTyrLeuLeuAspArgAspAspProSer 749
2212 TTTTTCAGCAGTAGAATTTGGAACATTTATTACAGCAGCCAGACGATCAG 2261
749 sPAlaLeuSerIysGlnLeuGlnProGlnValGlnGlyVal..... 762
2262 CAGCTACTACATCTTCTTGAAACGCTTAAGAGATGACAAATCTACT 2311
763 ...AspAsnIysMetSerGlnCysThrSerSerThrIleProSer 776
2312 GAACAGAAATGGAATGAGCAAAAGACAAATTTATATACCTCT 2336
seq_name: /cgn2_6/plodata/1/lna/6a_COMB.seq:us-08-915-213-1

```

```

seq_documentation_block:
; Sequence 1, Application US/08915213
; Patent No. 6020462
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,213
; FILING DATE: 20-AUG-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,473
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3736 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-915-213-1

alignment_scores:
Quality: 258.00 Length: 915
Ratio: 0.604 Gaps: 41
Percent Similarity: 46.667 Percent Identity: 19.781

alignment_block:
US-09-041-994-2 x US-08-915-213-1 ..
Align seg 1/1 to: US-08-915-213-1 from: 1 to: 3736
1 MetSerGlyLeuGlnGlyGlnAsnLeuAsp.....ProLeuAlaSerAspSe 15
|||
29 ATGGAAGGCGCGCGCGCGCGGACGACGACGAAAGAAATTAAGTTCTGCAAG 78
15 rArgIysArgIysLeuProCysAspThrProGlyGlnGlyLeuThrCys 32
|||
79 TCGAAAGAAAGAAAG.....TCTCGAG 98
32 erGlyGlnIysArgArgArgGlnGlnGlnIysSerIleGlnIleuLeu 48
|||
99 ATGCAAGCCAGATCTCGCGAGAGTAAAGATTCGAAATTTTATATGACGTT 148
49 AlaGlnLeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnVally 65
|||
149 GCTCAT.....CAGTTGCCACTTCCACATATATGATGATGCCA 186
65 sProAspIysCysAlaIleLeuIysGlnThrValArg.....GlnIleA 80
|||
187 TCTGTATTAAGCGCCTCTGTGATGAGGCTTACCATCATATTTGCTGTGA 236

```

[illegible]

```

1002 ATGACACCAAGAAATTCCTCAACACCACAGTCGATGTGTGTGATTAACGTT 1051
1003 GlnArgGluGlnAsnGlyThrArgProAsnProAsnProValGlyGlnG 386
1051 ..... 1051
386 yLeaArgProPromeValaGlyCysAsnSerSerValGlyMet... 401
1052 ..... GTGAGTGGTATTATTC 1067
402 ..... SerMetSerPro 405
1068 AGCAGCACTGTGATTTCTCCCTTCAACAACAGAAATGTGTCTTAAACCG 1117
406 AsnGlnGlyLeuGlnMetProSerSerArgValaGlyLeuAlaAspR 422
1118 GTTGAATCTTCAGATATGAAATGACTCAGCTATTCCACCAAGATTGAATC 1167
422 oSerThrThrGlyGlnMetSerGlyAlaArgTyrGlyGlySerSerAsnI 439
1168 AGAAGATACAAAGTAGGCTCTTTGACAAACTTAAGAAGAGACCTAGTGGTT 1217
439 leaLeaSerLeuThrProGlyIleProGlyMetGlnSerProSerSerTyrGln 455
1218 TAACCTTGTGGTGGCCCAAGCCGCTGGa..... GACACAAATCATTA 1255
456 AsnAsnAsnTyrGlyLeuAsnMetSerSer..... 465
1256 TCTTTAGATTATTTGGCAGCAGCAACAGAACTGATGACACAGCAACTTGA 1305
466 ..... ProProHisGlySerProGlyLeu... AlaProAsnGlnGln 478
1306 GGAAGTACCACTTATATATGATGATGATGCTCCCTCCACCCAAAGCAAAAT 1355
479 ..... AsnIleMetIleSerProAlaAsnArgGlySerProIlys 491
1356 TACAGAAATTAATTAATTGGCAATGTCTCA..... 1384
492 IleAlaSerHisGlnPheSerProValAlaGlyValHisSerPromeAl 508
1385 ..... TTACCACACCGCTGAACGCCAAAGCCACTTGC 1416
508 aSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerSerLeuSerA 525
1417 AAGTATGTCTGAC..... 1429
525 leaLeuGlnAlaIleSerGlnGlyValaGlyThrSerLeuLeuSerThrLeu 541
1430 ..... CCTGCACATCATCAAGAAAGTTGCATTAAATTA..... 1462
542 SerSerProGlyIleProIlysLeuAsnSerSerProAsnMetAsnIleThrGln 558
1463 ..... GAACCAAAATCCAGAG..... TCACGTGAACCTTTCTTTTACCAT 1500
558 nPro..... SerIlysv 562
1501 GCCCCAGATTTCAGAGATCAACAGCAACGATAGTCTCTCCATGGAAAGCACTAAC 1550
562 alSerAsnGlnAspSerIlysSerPro..... LeuGlyPheTyrCysAsp 576
1551 AAGGTCACCTGAGCCTATATATGATGCCAGAGATATATTGTTTATATGTCGAT 1600
577 GlnAsnProValaGluSerSerMetCysGlnSerAsnSerAlaGlyAspHisLe 593
1601 AGGAGATGTGTCAAATCAATTAAGATTGGAAATGTGTAGAAAAAATTTTTCG 1650
593 userAspIlyGluSerIlyGluSerSerValaGluGlyAlaGlnAsnGlnA 610
1651 TGAAGACACAGAGCAAG..... 1669
610 rglIleProLeuGlnSerIlysGlnIlyIysIlyLeuLeuGlnIleuLeuThr 626

```

```

1670 .. AACCATTTTCTACTGACGACACAGATTGAGTGTACT 1717
627 .....CysSerSerAsp.....ArgGlyHisSerSerLe 637
      |||||
1718 CCTATATCCAAATGAGATTACTTCCAGTTACCTTCTTCATCACT 1767
637 urthAnSerProLeuAspSerSer.....CysL 647
      |||||
1768 G.....TCACATTAGAAAGAGTGTCCGCAAGCCCTGAAAGCCAGTc 1811
647 ysgLuserSerValSerAlIthrSerProSerGlyValSerSerSer 663
      |||||
1812 CTCAAAGACCACTTACAGATTCCAGACAGACCAATACAAAGACTACT 1861
664 SerGlyGlyValSerSerThrSerAnethHisGlySerLeuLeuGlnI 680
      |||||
1862 GCTATGCCACCACTACCACTCCACCACTGATGATTAATAAACAAGTAC 1911
680 uLyHisArgGlyLeuHis..LysLeuLeuGlnAnsnGlyAnsnSerPro 696
      |||||
1912 AAAAAGCCGTATGGAAGACATTAAATGATGATTCATCTCCATCTCTTA 1961
696 IaGlyValAlaLysIleThrAlaGlnAlaThrGly.....LysAsp 709
      |||||
1962 CCCACATCATATAAAGAACTACTAGTCCACATCATCACCATATAGAGAT 2011
710 ThrSerSerIleThrSer.....CysGlyAspGly..... 719
      |||||
2012 ACTCAAGATCGGACAGCCCTCACCAACAGACGAGCAAAAGAGCATATGA 2061
719 ..... 719
2062 ACAGACAGAAAAATCTCATCCAAAGACCTTAACGTATTATCTGTGCTT 2111
720 .....AsnValValysGlnGlnLeuSerProLys..... 730
      |||||
2112 TGAGTCAAAGAACTACAGTCTCTGAGGAAGAACTAAATCCAAAGTACTA 2161
731 .....LysIysGlnAsnAlaLeuLe 738
      |||||
2162 GCTTTGACGAATGCTCAGAGAAAGCAAAATGGAACATGATGAGTTCCT 2211
738 u.....ArgTyrLeuLeuAspArgAspProSerA 749
      |||||
2212 TTTTCAAGCAGTAGAATTTGAACTTTATTACAGACGCCAGCATCATG 2261
749 sPaLAsuSerLysGluLeuGlnProGlnValGlnGlyVal..... 762
      |||||
2262 CAGCTACTACTACTCTTCTTGGAAACGTTAAAGGATGCAATCTACT 2311
763 ..AspAsnLysMetSerGlnCysThrSerSerThrIleProSer 776
      |||||
2312 GAACGAATGGAATGGAGCAAAAGCAATTTAATTAATACCTCT 2356
seq_name: /cgr2_6/ptodata/1/ina/6B_COMB.seq:US-09-148-547-1
seq_documentation_block:
  Sequence 1, Application US/09148547
  Patent NO. 6124131
  GENERAL INFORMATION:
  APPLICANT: Semenza, Gregg L.
  TITLE OR INVENTION: Hypoxia Inducible Factor-1 and Methods of Use
  FILE REFERENCE: 07265/151001
  CURRENT APPLICATION NUMBER: US/09/148,547
  CURRENT FILING DATE: 1998-08-25
  NUMBER OF SEQ ID NOS: 2
  SOFTWARE: Patentin Ver. 2.0
  SEQ ID NO 1
  LENGTH: 3736
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: CDS

```

```

LOC: 29..(2509)
US-09-148-547-1

alignment_scores:
    Quality: 258.00      Length: 915
    Ratio: 0.604        Gaps: 41
    Percent Similarity: 46.667    Percent Identity: 19.781

alignment_block:
US-09-041-994-2 x US-09-148-547-1 ..

Align seg 1/1 to: US-09-148-547-1 from: 1 to: 3736

1 MetSerGlyLeuGlyGluAsnLeuAsp.....ProLeuAlaSerAsp 15
   |||:||||| ||| ||| |||:||||:|
29 ATGAGAGGCGCCGCCGCCGACGACAAAGAAAAAGATAGTTCTGACG 78
   |||:||||| ||| ||| |||:||||:|
15 ratGlySerGlyLeuProCysAspThrProGlyGlnGlyLeuThrCys 32
   |||:||||| ||| ||| |||:||||:|
79 TCCAAAAAGAAAG.....TCCTCGAG 98
   |||:||||| ||| ||| |||:||||:|
32 ercGlyGluSerArgArgGlyGlnGlnGlnSerLysTyrIleGlyLeu 48
   |||:||||| ||| ||| |||:||||:|
99 ATCCAGCCACATCTCGCGGAGATGAAGATGTGAAGCTTTTATAGCTT 148
   |||:||||| ||| ||| |||:||||:|
49 AlaGluLeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnVal 65
   |||:||||| ||| ||| |||:||||:|
149 GCTCAT.....CAGTGCACCTTCACATATATGATGATCGCA 186
   |||:||||| ||| ||| |||:||||:|
65 sProAspLysCysAlaIleLeuLysGluThrValArg.....GlnIle 80
   |||:||||| ||| ||| |||:||||:|
187 TCTTGATTAAGCCTCTGTGATGAGCCTTACATCAGTATTGCGTGCA 236
   |||:||||| ||| ||| |||:||||:|
80 rglGlnIleLysGluGlnGlyLysThrIleSerAsnAspAspValGln 96
   ||:||||:| ||| ||| |||:||||:|
237 GGAACCTTCGAGCTGGT...CATTTGGATTGAAGATGACATG... 280
   ||:||||:| ||| ||| |||:||||:|
97 LysAlaAspValSerSerThrGlyGlnGlyValIleAspLysAspSer 113
   |||:||||| ||| ||| |||:||||:|
281 .....AAGCACAGAT 291
   |||:||||| ||| ||| |||:||||:|
113 uGlyProLeuLeuGlnAlaLeuAspGlyPheLeuPheValAlaAsn 130
   |||:||||| ||| ||| |||:||||:|
292 GAATTCTTTATTATGGAAGCCTGGAGGCTTTGTTGTTGCTTCACAG 341
   |||:||||| ||| ||| |||:||||:|
130 rglLysAlaAsnIleValPheValSerGluAsnValThrGlnTyrLeu 146
   |||:||||| ||| ||| |||:||||:|
342 ATGATGGTGACATGATTACTTTCTGTGAATGTGAACAAATACATGG 391
   |||:||||| ||| ||| |||:||||:|
147 TyrLysGlnGluAspLeuValAsnThrSerValTyrAsnIleLeuHis 163
   |||:||||| ||| ||| |||:||||:|
392 TTAACTCAGTTTGAACTAACAGCAGACAGTGTGTTGATTTTACATCC 441
   |||:||||| ||| ||| |||:||||:|
163 uGluAspArgLysAspPheLeuLysAsnLeuProLysSerThrValAsn 180
   |||:||||| ||| ||| |||:||||:|
442 ATGTGACCAATGAGGAATGAGAGAAATGCTTACACACAGA..... 485
   |||:||||| ||| ||| |||:||||:|
180 LysValSerThrAsnGlnProGlnArgGlnLysSer..... 192
   |||:||||| ||| ||| |||:||||:|
486 GCCTGTGAAAAAGGATAAGAACAAAAACACACACCGACCTTTTCTC 535
   |||:||||| ||| ||| |||:||||:|
193 .....HisThrPheAsnCys..... 197
   |||:||||| ||| ||| |||:||||:|
536 AGAATGATAGTGTACCTTAAGCCGAGGAAGAACTATGACATTAAGTC 585
   |||:||||| ||| ||| |||:||||:|
198 .....ArgMetLeuMetLysThrProHisAspIleLeuGluAspI 211
   |||:||||| ||| ||| |||:||||:|
586 TGCACATGAGAGGTATTGCACCTCAGCGCCACATTCACGTAATGAGA 635
   |||:||||| ||| ||| |||:||||:|
211 LeuAsnAlaSerProGluMetArgGlnArgTyrGluThrMetGlnCys 661
   |||:||||| ||| ||| |||:||||:|
636 CCAACAGTAACCAACT.....CAGTGTGG 661

```

```

228 AlaLeuSerGlnProArgAlaMetMetGluGluGlyGluAspLeuGln 244
    : : : : :
662 TATAAGAAACCACT.....ATGAC 681

244 rcysmetIleCysValAlaArgArgIleThrThrGluArgThrPhe 261
    : : : : :
682 CTGCTGGGTGGTGAATTTGTAACCCAT.....CCTACC 716

261 roSerAsnProGlu.....SerPheIleThrArgHis 271
    : : : : :
717 CATCAAAATTTGAAATTCCTTAGATAGCAAGACTTCTCTCAGTCGACAC 766

272 AspLeuSerGlyLysValAlaAsnIleAspThrAsnSerLeuArgSer 288
    : : : : :
767 ACCCTGGATATGAAATTTCTTATTGTGTGAA.....AGAAATTAC 807

288 rMetArgProGlyPhe.....GluAspIleIleArgArgCysIleGln 303
    : : : : :
808 CGAATTGATGGATATGAGCCAGAAAGACTTTTAGCCGCTCAATTTATG 857

303 rgPheSerLeuAsnAspGlyGlnSerTrpSerGlnLysArgHisTyr 319
    : : : : :
858 AATATTATCATGCTTTGACTCTGATCATCTGACCAAACTCATCAT... 904

320 GlnGluAlaTyrLeuAsnGlnHisAlaGluThrProValTyrArgPhe 336
    : : : : :
905 ...GATATGCTTACTAAAGACAGACACACAGACAGACAGATGCT 951

336 rLeuAlaAspGlyThrIleValThrIleGlnThrLysSerLysLeuPhe 353
    : : : : :
952 TGGCAAAAGAGGTGATATGCTCGGTGAAACTCAAGCACTGTCATAT 1001

353 rgAsnProValThrAsnAspArgHisGlyPheValSerThrHisPheLeu 369
    : : : : :
1002 ATAAACACCAAGATTCCTCAACCAAGTCATGTGATGTGAATTACGTT 1051

370 GlnArgGluGlnAsnGlyTyrArgProAsnProAsnProValGlyGln 386
1051 ..... 1051

386 ylleArgProPromeAlaGlyCysAsnSerSerValGlyGlyMet... 401
    : : : : :
1052 .....GTGAGTGGTATTATTC 1067

402 .....SerMetSerPro 405

1068 AGCAGCACTTGATTTTCTCCCTTCAACAAAGAAATGTCTCTTAACCG 1117

406 AsnGlnGlyLeuGlnMetProSerSerArgAlaTyrGlyLeuAlaAsp 422
    : : : : :
1118 GTTGAATCTTCAGATATGAAATGACTCAGTATTCACCAAGTTGAATC 1167

422 oSerThrThrGlyGlnMetSerGlyAlaArgTyrGlyGlySerSerAsn 439
    : : : : :
1168 AAGAAGTACAACTAGCCCTTTGACAAACTTAAGAAGAACTGATGCTT 1217

439 lAlaSerLeuThrProGlyProGlyMetGlnSerProSerSerTyrGln 455
    : : : : :
1218 TTACTTTGCTGGCCCGCCGCTGCA.....GACACAATCATATA 1255

456 AsnAsnAsnTyrGlyLeuAsnMetSerSer..... 465
    : : : : :
1256 TCTTTAGATTTTGGCAGACAGACAGAAACTGATGACCAGCAACTTGA 1305

466 .....ProProHisGlySerProGlyLeu...AlaProAsnGlnGln 478
    : : : : :
1306 GGAAGTACCATATATATGATGATATGCTCCCTCACCAGCAAGAAAT 1355

479 .....AsnIleMetIleSerProArgAsnArgGlySerProLys 491
    : : : : :
1356 TACAGAAATATAATTTGGCAATGTCTCA..... 1384

```

```

492 IleAlaSerHisGlnPheSerProValAlaGlyValHisSerProMetAl 508
    : : : : :
1385 .....TTACCCACCGCTGAAGCCCAAGCCACTTCG 1416

508 aSerSerGlyAsnThrGlnGlyAsnHisSerPheSerSerSerLeuSer 525
    : : : : :
1417 AAGTATGCTGAC..... 1429

525 lAlaGlnAlaIleSerGluGlyValGlyThrSerLeuLeuSerThrLeu 541
    : : : : :
1430 ....CCTCCACTCAATCAGAAAGTTGCATTAAATTA..... 1462

542 SerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleThrG 558
    : : : : :
1463 ...GAACCAATCCAGAG.....TCACGTGAACATTCTTTTACCAT 1500

558 nPro.....SerLysVal 562
    : : : : :
1501 GCCCAGATTCAGATCAGACACCTAGTCTTCGATGGAGCACCTAGAC 1550

562 aLysAsnGlnAspSerLysSerPro.....LeuGlyPheTyrCysAsp 576
    : : : : :
1551 AAGATTACCTGAGCGCTAATAGTCCAGTGAATATTGTTTATGTGAT 1600

577 GlnAsnProValGlnSerSerMetCysGlnSerAsnSerArgAspHisLe 593
    : : : : :
1601 AGTGATATGGTCATGAATGAATTCAAGTTGGAAATGTAGAAAACTTTTGC 1650

593 uSerAspLysGlnSerLysGlnSerSerValGluGlyAlaGluAsnGln 610
    : : : : :
1651 TGAAGCACAGAAAGCAAG..... 1669

610 rgGlyProLeuGlnuSerLysGlnHisLysLysLeuGlnLeuLeuThr 626
    : : : : :
1670 ..AACCATTTCTACTCGACAGACAGATTTAGACTGTGAGATGTTAGCT 1717

627 .....CysSerSerAspAsp.....ArgGlyHisSerSerLe 637
    : : : : :
1718 CCTATATCCCAATGATGATGACTTCCAGTTACGTTCTCTTGATCAGCT 1767

637 uThrAsnSerProLeuAspSerSer.....CysL 647
    : : : : :
1768 G.....TCACCATTAAGAAGCAGTTCGCCAAGCCCTGAAGCGCAAGTC 1811

647 ysgLysSerSerValSerValThrSerProSerGlyValSerSerSerThr 663
    : : : : :
1812 CTCMAAGCACAGTTACAGTATTCAGCAGACACTCAATACAAAGACCTACT 1861

664 SerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeuGln 680
    : : : : :
1862 GCTAATGCCACCACTACCTCCACTGCCACTGATGAATTAATAACAGTGAC 1911

680 uLysHisArgIleLeuHis...LysLeuLeuGlnAsnGlyAsnSerPro 696
    : : : : :
1912 AAAAGACGCTATGGAAGCATTAATAATATGATTCGATCTCATCTCCGA 1961

696 lAlaGluValAlaLysIleThrAlaGlnAlaThrGly.....LysAsp 709
    : : : : :
1962 CCCACATACATTAAGAAGAACTACTAGTCCACATCATCAACATATAGAGAT 2011

710 ThrSerSerIleThrSer.....CysGlyAspArgLys..... 719
    : : : : :
2012 ACTCAAAAGTCGACAGCCCTCACCAAAACAGAGGAGAAAAGAGTATAGA 2061

719 ..... 719

2062 ACAGACAGAAAATGTCATCCAGAAGCCGTAAAGCTGTTATCTGCGCT 2111

720 .....AsnValValLysGlnGluGlnLeuSerProLys..... 730
    : : : : :
2112 TGAGTCAAGAACTACAGTCTCTGAGAGAAAGAACTAAATCCAAAGATACTA 2161

731 .....LysLysGluAsnAsnAlaLeuLe 738

```



```
2162 GCTTGCACATGCTCAGAGAAAGCAAAATGAAACATGATGTTCACT 2211
738 U.....ArgTyleuLeuAspArgAspProSera 749
2212 TTTTCAGACGAGTAGAATTTGAAACATTATTACAGCCAGCAATCANG 2261
749 sPAlaLeuSerIysGluLeuGlnProGlnValGluGlyVal..... 762
2262 CAGCTACTACTACCTTCTTGTGAAACGCTAAAGATGCAAAATCTACT 2311
763 ...AspAsnIysMetSerGlnCysThrSerSerThrIleProSer 776
2312 GAACAGAAATGAAATGAGCAAAAGCAATATTATTATACCTCT 2356
seq_name: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq: PCT-US96-10251-1

seq_documentation_block:
: Sequence 1, Application PC/PTUS9610251
: GENERAL INFORMATION:
: APPLICANT: The Johns Hopkins University School of Medicine
: TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: USA
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US96/10251
: FILING DATE: 06-JUN-1996
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Halle, Lisa A.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: 07265/053W01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619/678-5070
: TELEFAX: 619/678-5099
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3736 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
: PCT-US96-10251-1

alignment_scores:
  Quality: 258.00      Length: 915
  Ratio: 0.604        Gaps: 41
  Percent Similarity: 46.667      Percent Identity: 19.781

alignment_block:
US-09-041-994-2 x PCT-US96-10251-1 ..
Align seg 1/1 to: PCT-US96-10251-1 from: 1 to: 3736

1 MetSerGlyLeuGluAsnLeuAsp.....ProLeuAlaSerAsp 15
29 ATGAGAGGCGCGCGCGCGAGCAAGCAAAAGAAATATTGTAACG 78
15 rArgIysArgIysLeuProCysAspThrProGlyGlnGlyLeuThrCys 32
79 TCGAAGAGAGAAAG.....TCTCGAG 98
```

```
32 erGlyIuLysArgArgArgGluGlnGlnSerIysThrIleGluLeu 48
99 ATGACCCGCAATCTCGCGAGTAAGAAATCTGAAGTTTATTATGACT 148
49 AlaGluLeuIleSerAlaAsnLeuSerAspIleAspAsnPhenAlly 65
149 GCTCAT.....CAGTTGCCACTTCCACATTAATGATAGTTCCGA 186
65 sProAspIysCysAlaIleLeuLysGluThrValArg.....GlnIleA 80
187 TCTGTATAGGCGCTCTGTATGATGAGGCTTACATCAGCTATTGCGGTGA 236
80 rGlnIleLysGluGlnGlyIysThrIleSerAsnAspAspAlaGln 96
237 GGAAACTTCTGATGCTGCT...GATTGTGATATTCAGATACATG... 280
97 LysAlaAspValSerSerThrGlyGlnGlyValIleAspLysAspSer 113
281 .....AAGCACAGAT 291
113 uGlyProLeuLeuGlnAlaLeuAspGlyPheLeuPheValAlaAsn 130
292 GAATTCCTTTTATTGAAAGCCTTGATGCTTTGTTATGCTTCACAG 341
130 rGluAlaAsnIleValPheValSerGluAsnValThrGlnIleGln 146
342 ATGATGCTGATGATTTTACATTTCTGATTAAGTGAACAAATACAGGA 391
147 TyrLysGlnGluAspLeuValAsnThrSerValTyrAsnIleLeuHis 163
392 TTAACCTCACTTGAACCTAGACAGACAGTGTGTTGATTATTCATCC 441
163 uGluAspArgLysAspPheLeuLysAsnLeuProLysSerThrValAsn 180
442 ATGTGACCATGAGAAATGAGAAATGCTTACACACAGA.....AATG 485
180 LysValSerThrAsnGluProGlnArgGlnLysSer..... 192
486 GCCTTGTGAAAAAGGTAAAGAACAAACACAGGAGGAGCTTTTCTTC 535
193 .....HisThrPheAsnCys..... 197
536 AGAATGAAGTACCTTACCTAGCCGAGAGAACTATTAACATTAAGTC 585
198 .....ArgMetLeuMetLysThrProHisAspIleLeuGluAspI 211
586 TGCACATGGAAGGTATTCACCTGCACAGCCACCATTCACGTATATGATA 635
211 LeuAsnAlaSerProGluMetArgGlnArgTyrGluThrMetCysPhe 227
636 CCAACAGTAACCACT.....CAGTGTGG 661
228 AlaLeuSerGlnProArgAlaMetMetGluGluGlyAspLeuGlnSe 244
662 TATTAAGAAACCACT.....ATGAC 681
244 rCysMetIleCysValAlaArgArgIleThrThrGlyGluArgThrPhe 261
682 CTGCTTGCTGCTGATTTTGTGAACCATT.....CCTCAC 716
261 rSerAsnProGlu.....SerPheIleThrArgHis 271
717 CATCAAAATATGAAATTCCTTAGATAGCAACACTTCTCAGTCAGAC 766
272 AspLeuSerGlyLysValValAsnIleAspThrAsnSerLeuArgSer 288
767 AGCTTGATATGAATTTTCTATTGTGTAGAA.....AGAAATTC 807
288 rMetArgProGlyPhe.....GluAspIleIleArgArgCysIleGln 303
808 CGAATTGATGGATATGAGCCAGAGAAAGACTTTAGCCGCTCAATTATG 857
```



```

1  APPLICANT: Montminy, Marc R.
2  TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF
3  TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND MITOGEN
4  TITLE OF INVENTION: RESPONSIVE GENES
5  NUMBER OF SEQUENCES: 3
6  CORRESPONDENCE ADDRESS:
7  ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
8  STREET: 444 South Flower Street, Suite 2000
9  CITY: Los Angeles
10 STATE: California
11 COUNTRY: USA
12 ZIP: 90071
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: Floppy disk
15 COMPUTER: IBM PC compatible
16 OPERATING SYSTEM: PC-DOS/MS-DOS
17 SOFTWARE: PatentIn Release #1.0, Version #1.25
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/08/194,468
20 FILING DATE: 10-FEB-1994
21 CLASSIFICATION: 435
22 ATTORNEY/AGENT INFORMATION:
23 NAME: Reiter, Stephen E.
24 REGISTRATION NUMBER: 31,192
25 REFERENCE/DOCKET NUMBER: P41 9672
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: (619)-546-4737
28 TELEFAX: (619)-546-9392
29 INFORMATION FOR SEQ ID NO: 1:
30 SEQUENCE CHARACTERISTICS:
31 LENGTH: 7326 base pairs
32 TYPE: nucleic acid
33 STRANDEDNESS: single
34 TOPOLOGY: linear
35 MOLECULE TYPE: DNA (genomic)
36 HYPOTHETICAL: NO
37 ANTI-SENSE: NO
38 FEATURE:
39 NAME/KEY: CDS
40 LOCATION: 1..7323
41 OS-08-194-468-1

```

[illegible]

```
alignment_block:
US-09-041-994-2 x US-08-194-468-1  . .
```

Align seg 1/1 to: US-08-194-468-1 from: 1 to: 7326

[illegible]

940	IQPRLVGLYAspIyrAsnThrSerLeuProArgProAlaLeuGlyLys	956
5813	ASGCT.....ACCAACGAGGGCCAGCTCCCCACCCCTGCC	5850
957	SerIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArg	973
5851	CAGCCCCACCTGCA.....GCAGTAGAAGAGCGCGGCA	5885
973	oVal.....LeuGlnGlnGlnGlnMetLeuGln.....	983
5886	AATGACGTGAGGGCCGACGACGACACCTATACCAGCAACATCA	5935
984MetArgProGlyuIleProMetGly.....	992
5936	ACAATGCGATGGCCCCAGAGCGTGCAGGTATGGGACCCAGAACCA	5985
993MetGlyAlaAsn.....ProTyrGlyLys	1000
5986	ATGACTCCCTGGGCGCTCAATGTGCTCCCGTCCCAACCAATGCACTGGCC	6035
6036	TGTCAATGTAGTATGACCACTGGCAGTGGCAGACAGCAACCATCTCTC	6085
1017	eGlnGlnValSerIleGlyThrGlnAsnArgProLeuLeuAlaGlnSer	1033
6086	ACACGACCGCGATGCCAGG...ATGCCCAAGGCTGTAAATGTCATCAG	6132
1034	LeuAspArgLeuValGlyProProSerAsnLeuGlnGlnSerAspGly	1050
6133	GCCCGAGCAGAGTGGTGGGCGACGGATGCCCAATGTGCAAGCCAAAG	6182
1050	u.....ArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerA	1063
6183	GAGCATCTGCGCAAGTGGCTGCAAGACCTGTGCGACCTTAATGATCAC	6232
1063	snThrAspAlaThrGlyLeuGlnGluIuIleAspArgAlaLeu...GlyTle	1078
6233	CCAGCTCTCTCAGACGACGACGAGCGTGCATCCCTAAATGCAAC	6282
1079	ProGluLeuValAsnGlnGlyValAlaLeuGluIuProLysGlnAspAlaPh	1095
6283	CCACAGCTATG.....GCAGCTTT	6302
1095	eGlnGlnGlnAlaAlaValMetMetAspGlnLysAlaGlyLeuTyrG	1112
6303	CATCAACACGCGCACAGCGCAAGTATGTGCCAATACAGCTGGCATGACG	6352
1112	Iy.....GlnThrTyrProAlaGlnGlyProProMetGlnGly	1124
6353	CCGAGCGCGAATTCAATCCGACGCTGTATGACAGCC.....CAGCCT	6396
1125	GlyPheHisLeuGlnGlyGlnSerProSerPheAsnSerMetAsnGln	1141
6397	GGCATGACACCAAGCTATGTTGCAAAACCTGAACGCAATGCAAGCTGG	6446
1141	metAsnGlnGlnGlyAsnPheProLeuGln.....GlyMetH	1154
6447	TGTGGCAGGGCTGTGTGGCTCCACCAACCAACGACATGGGAGGCTGA	6496
1154	isProArgAla.....AsnIleMetArgProArgThrAsn.....	1165
6497	ATCCCGAGGACAACTGTGAACATCATGAACCCAGACACACCCCAAC	6546
1166ThrProLysGlnLeuArgMetGlnLeuGlnGlnArgLe	1178
6547	ATGACAAACATGAATCCACAGTACCGAATAATGTGAGGAGACAGCTGT	6596
1178	uGlnGlnGlnGlnPheLeuAsnGlnSerArgGlnAlaLeuGlnLeuTyrSm	1195
6597	ACAGCACACGACGACGACGACGACAAACACACGACGACGACGACAAAC	6646
1195	eGlnAsnProThr...AlaGlyGlyAlaAlaValMetArgProMetMet	1210

```

6647 AAATATGTCACGCTTGGCGGGGGGSCATGGCGGACACAGCCAGTCTCCAG 6696
1211 GlnProGln.....GlnGlyPheLeuAsnAlaGlnMetValAlaGln 1225
6697 CAGCCCAAGAGACCTGGAGGTATATGCCCAAGCATGCAGACAGAACGCAT 6746
1225 gSerAtgGluLeuLeuSerHisHisPheArgGlnGlnArgValAlaMet 1242
6747 GCAAAGCAGCTCCCATCCAGGCGAGCTCCAGGCGCAGATGGCTGCTC 6796
1242 eMetGlnGlnGlnGlnGlnGlnGln..... 1252
6797 CAATGGAGACAATGGCCAGATGGGCGAGCCTGGCTAGGGGCGACACGC 6846
1253 .....GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1265
6847 ACCCCCAATATCCAGCAGGCGCCGCGACAAAGCATCTCTCAGACAGACGA 6896
1265 nGlnGlnThrGlnAlaPheSerProProProAsnValThrAlaSerPro 1282
6897 GATGAAGCAACAATTTGGGTACACAGCCAGCCGAACCCCATGAGCCGCC 6946
1282 eMetAspGlyLeuAlaGlyProThrMetProGlnAlaProProGln 1298
6947 AGCAGCACAATGCTCTCAGGACAGCCAGCGCTCACACATCTCCTGGCCAG 6996
1299 GlnPheProGlyGlnProAsnTyrglyMetGly..... 1310
6997 CAGATGCCACATCCCTTAGTAACAGAGTGGCATCTCCAGCCCTGTGCA 7046
1310 nGlnProAspProAlaPheGlyArgValSerSerProProAsnAlaMet 1327
7047 GCTCTCAGCGCCCA.....TCCCAACCTCCACATCTCCAGCC 7084
1327 eSerSerArgMetGlyProSerGlnAsnProMetMetGlnHisProGln 1343
7085 CGTCACACAGGATACAAACCCAGCTTCACACACCATGTTCACCCCG 7134
1344 AlaAlaSerIleTyrglnSerSerGluMetLysIleTyrglyProSerGly 1360
7135 ACT.....GGAACCCCTCACACCTCG 7154
1360 nLeuAlaArgAsnSerSerPheSerGlnGlnPheAlaHisGlnGly 1377
7155 ACTCGAGTCACCATGGCCAGCTCCATGATCAG...GGACACCTGGGA 7201
1377 snProAlaValTySerMetVal 1384
7202 ACCCTGAACAGAGTGCATGCTC 7224

seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:US-08-961-739-1
seq_documentation_block:
: Sequence 1, Application US/08961739A
: Patent No. 6063583
: GENERAL INFORMATION:
: APPLICANT: Montminy, Marc R.
: TITLE OF INVENTION: Methods for Treating Diabetes Mellitus
: FILE REFERENCE: SALKI1650-1
: CURRENT APPLICATION NUMBER: US/08/961,739A
: EARLIER FILING DATE: 1997-10-31
: EARLIER APPLICATION NUMBER: US 194,468
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 7326
: TYPE: DNA
: ORGANISM: Mus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(7326)
```

```

: NAME/KEY: misc_feature
: LOCATION: (1)...(7326)
: OTHER INFORMATION: n = A,T,C or G
US-08-961-739-1

alignment_scores:
  Quality: 254.50      Length: 591
  Ratio: 0.887         Gaps: 27
  Percent Similarity: 48.562  Percent Identity: 24.873

alignment_block:
US-09-041-994-2 x US-08-961-739-1 ..

Align seg 1/1 to: US-08-961-739-1 from: 1 to: 7344

857 SerLeuAspSerProValSerValGlySerSerProProValLysAsn1 873
||||| |||||||:|||||: |||:|||||:
5644 AGTTGCTTCTTCTTACCTAGCACCCGGGACTCTACACAGCAG.. 5691
873 eSerAlaPheProMetLeuProLysGlnProMetLeuGlyLysPro 890
5692 .....CCAGCAGACCCCAACACACAGCCGCCAGCCAGCTTC 5731
890 rgMetMetAspSerGlnGluAsnTyrglySerSerMetGlyLysPro 906
5732 AG.....CCTTCA 5739
907 ArgAsnValThrValThrGlnThrProSerSerGlyAspTyrGlyLeu 923
5740 .....CCTGTACATGTCACAGCA.....GCTTCCC 5768
923 oAsnSerLysAlaGlyArgMetGluProMetAsnSerSerMetGly 940
5769 TAAT.....GTAGCCGAGCTCAGCCCAACAAATAGTGTCTGGGGA 5812
940 rgProGlyLysPtyrAsnThrSerLeuProArgProAlaLeuGly 956
5813 AGCCT.....ACCAACAGGTGTCAGCTCCGCCACCCCTGCC 5850
957 SerIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArg 973
5851 CAGCCCCACCTGCA.....GCAGTGAAGCAGCCCGGCA 5885
973 oVal.....LeuGlnGlnGlnGlnMetLeuGln..... 983
5886 AATTGAAGCTGAGGCCAGCAGCAGACGACCATATACGAGAACATCA 5935
984 .....MetArgProGlyGluIleProMetGly..... 992
5936 ACATGGCATGCCCCAGCAGCTGACGTATGGGACCCAGAACCA 5985
993 .....MetGlyAlaAsn.....ProTyrGly 1000
5986 ATGACTCTGTGGGCTGGAATGTGCCCGTCCCAACCAAGTCAGTGGGC 6035
1000 nAlaAlaAlaSerAsnGlnLeuGlySerTyrProAspGlyMetLeuSer 1017
6036 TGTCAATGTCTAGTATGCCACCTGGGAGTGGGAGCAGCAGCAACCACTCC 6085
1017 eGluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSer 1033
6086 ACACAGAGCCGATGCCAGC...ATGCCAGGCGGTATATGTCATGAC 6132
1034 LeuAspAspLeuValGlyProProSerAsnLeuGlnGlyGlnSerAsp 1050
6133 GCCCAGGACAGAGTGGCTGGCCAGCATGCCCAATGTGCACCAACAG 6182
1050 u.....ArgAlaLeuAspGlnLeuHisThrLeuLeuSer 1063
6183 GAGCATCTGCCCAAGTCCCTGCACAGACCTGTACGAGACCTAAAGTCA 6232
1063 snThrAspAlaThrGlyLeuGlnGluIleAspArgAlaLeu...GlyIle 1078
```

```

6233 CCAAGCTCTCTCAGCAGCAGCAGCTGCTGAACATCTTAATATCAAC 6282
1079 ProGUleuValasnGlnGlnAlaLeuGluProLysGlnAspAlaPh 1095
6283 CCACAGCTAATG.....GCAGCTTT 6302
1095 eGlnGlnGlnAlaAlaValaMetMetAspGlnLysAlaGlyLeuTyrG 1112
6303 CATCAAAACAGCGCAGCAGCAAGATATGTGGCAATCAGCAGCTGCATC 6352
1112 Ly.....GlnThrTyrProAlaGlnGlnGlyProPrometGlnGly 1124
6353 CCCAGCCCGGAGCTTCAATCCAGCCTGTGATGCAAGCC.....CAGCCT 6396
1125 GlyPheHisLeuGlnGlnGlnSerProSerPheAsnSerMetMetasnG 1141
6397 GGCATGCAACGAGCAGCTAGTTGCAAAACCTGACAGCAATGCAAGCTG 6446
1141 nMetasnGlnGlnGlyAsnPheProLeuGln.....GlyMetH 1154
6447 TGTGCCAGCGCTGTGTGCTCTCCACCACACAGCAATGGAGGCTGTA 6496
1154 IsPProArgAla.....AsnIleMetArgProArgThrAsn..... 1165
6497 ATCCCAAGGAGCAGAGCTGTAACATCATGACAGCAGACCAACCCCAAC 6546
1166 .....ThrProLysGlnLeuArgMetGlnLeuGlnGlnArgLe 1178
6547 ATGACCAACATGTAATCCACAGTACCGAATAATGGTAGAGGACAGCTG 6596
1178 uGlnGlnGlnGlnPheLeuasnGlnSerArgGlnAlaLeuGlnLeuLys 1195
6597 ACAGACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAC 6646
1195 eGlnLysAsnProThr...AlaGlyAlaAlaValaMetArgPrometMet 1210
6647 AAAATGATGCCAGCTTGCGCGGGGAGATGGCGGACAGCAGCCAGTTCAG 6696
1211 GlnProGln.....GlnGlyPheLeuasnAlaGlnMetValaGlnInr 1225
6697 CAGCCCAAGCAGCAGCTGAGGTATGCCCAAGCCATGACAGCAGCAAGC 6746
1225 gSerArgGlnLeuLeuSerHisIshPheArgGlnGlnArgValaIleMet 1242
6747 GCACAGCAGCAGCTCCCATCCAGGCGAGCTCCATGAGCGCCAGATGCTG 6796
1242 eMetGlnGlnGlnGlnGlnGlnGln..... 1252
6797 CAATGGGACAACTTGGCAGATGGCGAGCTGGCTAGGGGAGCAGACAG 6846
1253 .....GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1265
6847 ACCCCATATATCCAGCAGCGCCCTGACCAAGGATTCCTGACAGCAGCA 6896
1265 nGlnGlnThrGlnAlaPheSerProProAsnValThrAlaSerProS 1282
6897 GATGAAGCAACAATTTGGCTACCCAGCAGCCGGAACCCCATGAGCCGCC 6946
1282 eMetAspGlyLeuAlaGlyProThrMetProGlnAlaProGln 1298
6947 AGCAGCAGCATGCTCTCAGCAGCAGCAGCAGCCTCATCATCTCCCTGG 6996
1299 GlnPheProTyrGlnProAsnTyrGlyMetGly.....G 1310
6997 CAATATGCCACATCCCTTAGTAACAGGTGCGATCTCCAGCCCTGTGTA 7046
1310 nGlnProAspProAlaPheGlyArgValaSerSerProProAsnAlaMet 1327
7047 GTTTCACAGGCCCCAA.....TCCCAACTCCACATTCACAGCC 7084
1327 eSerSerArgMetGlyProSerGlnAsnProMetMetGlnHisProGln 1343

```

```

7085 CGTCACACAGGATACACCCAGCCTTCACACACACCATGTTTACCCCGAG 7134
1344 AlaAlaSerIleTyrGlnSerSerGlnMetLysGlyTyrProSerGlyAs 1360
7135 ACT.....GGAACCCCTCACCCCTGG 7154
1360 nLeuAlaArgAsnSerSerPheSerGlnGlnGlnPheAlaHisGlnGlyA 1377
7155 ACTCGCAGTCCACCATGGCCAGCTCCATGGATGACG...GGACACCTGGGA 7201
1377 snProAlaValTyrSerMetVal 1384
7202 ACCCTGACACAGATGCATGCTC 7224

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-045-806-3
seq_documentation_block:
Sequence 3, Application US/08045806
Patent No. 5378822
GENERAL INFORMATION:
APPLICANT: Bradford, Christopher Alan
APPLICANT: Dolwick, Kristin Marie
APPLICANT: Poland, Alan
TITLE OF INVENTION: Ah Receptor cDNA and Method of
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Tilton, Fallon, Lungmus & Chestnut
STREET: 100 South Wacker Drive, Suite 960
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-4002
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/045, 806
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rentress, Susan B.
REGISTRATION NUMBER: 31,327
REFERENCE/DOCKET NUMBER: NU-9207
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312)-456-8000
TELEFAX: (312)-456-7776
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5261 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 383..2927
US-08-045-806-3

alignment_scores:
Quality: 252.00 Length: 1051
Ratio: 0.545 Gaps: 43
Percent Similarity: 43.958 Percent Identity: 19.696

alignment_block:
US-09-041-994-2 x US-08-045-806-3 ..
Align seg 1/1 to: US-08-045-806-3 from: 1 to: 5261
15 SerArgLysArgLysLeuProCysAsp.....ThrProGln 26

```

```

416 AGTGGAGGCGGAAAGCGGTGACAGAAAACTAAAGCCATCCAGC 465
26 yGlnGlyLeuThrCysSerGlyGlyLysArgArgGlnGlnSerL 43
466 TGAAGAAATCAATCAATCTCTCCAAAGCGGCATAGAGCCGACTTATA 515
43 yStrYlIeGlnGlnLeuAlaGluLeuIleSerAlaAsnLeuSerAspIle 59
516 CAGAGTTGACCGCTTGTGCTGCTGCTGCTTCCCA...CAAGAGTGT 562
60 AspaSnPheAsnValLysProAspLysCysAlaIleLeuLysGluThrIva 76
563 ATTAAT...AAGTTGACAAACTTTTCAGTCTTAGGCTCAGCGCT 603
76 IArgGlnIleArg...GlnIleLysGlnGln 86
604 CAGTTACTGAGAGCCAAAGACCTTTGATGATGATTAATAATCCGCC 653
86 LysThrIleSerAsnAspAspValGlnLysAlaAspValSerSer 102
654 CTACTGAAGAAACGAGCGGAGGATTAAGTGAAGCAGCAATTTTGA 703
103 ThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeuLeuLeuG 119
704 GAAGGCGCTGAACCTTACAAGAA...GGAGAAATTTCTATTATACA 741
119 nAlaLeuAspGlyPheLeuPheValValAsnArgGlnAlaAsnIleValP 136
742 GCGCTGCAATGGCTTTGATTAAGTGTACACACAGATGCTTTGCTTTT 791
136 heValSerGlnAsnValThrGlnThrLeuGlnThrLysGlnGlnAspLeu 152
792 ATGCTTCTTACTACTACAAAGATATATCTAGGGTTTTCAGACTGTGATGC 841
153 ValAsnThrSerValLysAsnIleLeuIleGlnGlnLysAspArgLysAsp 169
842 ATACATCAGAGTGTATAGAACTTATCCATACCAAGACGAGCTGATTT 891
169 LeuLysAsnLeuProLysSerThrValAsnGlyValSerThrAsnG 186
892 T... 892
186 LurProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeuMetLys 202
893 ...CAGCGTACGCTACACTGGGCATTAAAT... 919
203 ThrProHisAspIleLeuGlnAspIleAsnAlaSerProGlnMetArgG 219
920 ...CTTTCTCAGTGTACAGAGTCTGACAAAGAAATTGAAGAGCCACTGG 966
219 nArgTyrGlnThrMetGlnCysPheAlaLeuSerGlnProArgAlaMetL 236
967 TCCTCCCGACAGTACGTCTGTAT...AACCCACAGACAGATTC 1007
236 etGlnGlnGlnGlnAspLeu...GlnSerCysMetIleCysValAlaArg 251
1008 CTCGAGAAATCTCTCTTATGAGAGGTCTTCATATATGTCGTCTAAG 1057
252 ArgIleThrThrGlnGlnArgThrPhe... 260
1058 TGTCTGCTGATATATCATCTGTTTCTGCGATGAATTTCCAAAGGAA 1107
260 ... 260
1108 GTTAAAGTATCTTCATGAGCAAGAAAGAAAGGAAGATGATCAATAC 1157
261 ...ProSerAsnProGlu 265
1158 TTCACCTCAGATGGCTTTGTTGCGATAGTACTACCTCAGCCACCA 1207
266 Ser...PheIleThrArgHis.. 271

```

```

1208 TCCTACTTGAAATCCGAGCCAAATTTTATCTTTGAAACCAACAAACA 1257
272 ...AspLeuSerGlyLysValValAsnI 280
1258 ACTAGACTTACACCTTATGTTGTGTATGCCAAAGAAATGTTTATAG 1307
280 IeAspThrAsnSerLeuArgSerSerMetLarProGlyPheGlnAspIle 296
1308 GATTAATGAAAGACAGCTGTGACAGAGAGCTTCAGTATTCAGTTAT 1357
297 ...IleArg... 298
1358 CATCAGCTGATATGCTTTATGTCGCGAGTCCCATATCCGAATGATTA 1407
299 ...ArgCysIleGlnArgPhePheSerLeuAsnAspG 310
1408 GACTGAGAAAGTGCATGATAGTTTCGGGCTCTTACAAAAAACAAC 1457
310 LysIleSerThrPheSerGlnLysArgHisTyrGlnGlnAlaTyrLeuAsnGly 326
1458 GATGAGCTTGGTCCAGTCT...AATGACGCGCTGTATTAAAAAATGGA 1504
327 HisAlaGluThrProValTyrArgPheSerLeuAlaAspGlyThrIleVa 343
1505 AGA...CCAGATTATATCATTTGT 1524
343 LThrAlaGlnThrLysSerLysLeuPheArgAsnProValThrAsnAspA 360
1525 AACT...CAGAGACCACTAAACAGATGAGG 1550
360 rGHisGlyPheValSerThrHisPheLeuGlnArgGlnGlnAsnGlyTyr 376
1551 AA...CGAAAACGAATAATCAGAG 1585
377 ArgProAsnProAsnProValGlyGlnGlyLysLeuArgProProMetAlaG 393
1586 TTGGCTTTTATGTTTACCAGTGGAAAGCTGTGTGTATGAGCCAAACAA 1635
393 yCysAsnSerSerValGlyGlyMetSerMetSerProAsnGlnLeuG 410
1636 CCCTTTCTGTCATTA...ATGATATCC...TTAC 1664
410 LmetProSerSerArgAlaTyrGlyLeuAlaAspProSerThrThrGly 426
1665 CACTAAGACTTAAATAATGCGACTAGTGAAGAAAGCTGTGACACATCC 1714
427 GlnMetSerGlyAlaArgTyrGlyLysSerSerAsnIleAsnSer... 441
1715 ACTTAAGCAAGAGCTGTCAATCTTAAGTTCCCTGCTGCTGCCATGAT 1764
442 ...LeuThrProGlyProGlyMetGlnSerP 451
1765 GCAACAGATGAGTCTATTTATCTCTATCCGCTTCAAGTACTCAAGTA 1814
451 roSerSerTyrGlnAsnAsnAsnTyrGlyLeuAsnMetSerSerProPro 467
1815 CTCGACCTTTTGAATAACCTTTTCAAGCAATCTATGAATGAATGCAGA 1864
468 HisGlySerProGlyLeuAlaProAsnGlnGlnAsnIleMetIleSerPr 484
1865 AATTGGCAAGATTAATGACCG... 1888
484 oArgAsnArgGlySerProLysIleAlaSerHisGlnPheSerProValA 501
1889 ...ATGGAAATGATACTACTCTGAACATGAG... 1918
501 IagLysValHisSerProMetAlaSerSerGlyAsnThrGlnAsnHisSer 517
1919 ...CAAATGACACGCTCAGAGATGGAATCATTTGCTGAGGCTCACCCA 1966
518 ...PheSerSerSerLeuSerAlaLeuGlnAlaIleSerGlnG 532
1967 GGGCTTTTCAAGATAGTAAACAGATGACTTGTACAGCATATGAAGAA 2016

```

```

532  ValGlyThrLeuLeuSerThrLeuSerSerProGlyProlLysLeuA 549
      :::::|:::|:::|:::|
2017 CCTAGCGATTGATTTT..... 2032
549  spAsnSerProAsnMetAsnIleThrGlnProSerLysValSerAsnGln 565
      :::::|:::|:::|:::|
2033 .....GAAGACATCAGACATCAGACATGAAATTTTTCAGAAAT 2074
566  AspSerLysSerProLeuGlyPheTyrCysAspGlnAsnProValGluSe 582
      |||:::|:::|:::|
2075 GATTTTTCGTGAGTTGACTTC..... 2098
582  rSerMetCysGlnSerAsnSerArgasp...HisLeuSerAspLysGluS 598
      |||||:::|:::|:::|
2099 .....AGAGCATTTGACTTAAAGCAT..... 2119
598  eLysGlnSerSerValGluGlyAlaGlnAsnGlnArgGlyProlLeuGln 614
2119 ..... 2119
615  SerLysGlyHisLysLysLeuLeuGlnLeuThrCysSerSerAspAs 631
      :::::|:::|:::|:::|
2120 .....GAATCTCTGACGTATGTCACAGAT.. 2143
631  pArgGlyHisSerSerLeuThrAsnSerPro..... 641
      |||||:::|:::|:::|
2144 .....TCATTAAAGTCTCCCTTCATACCTTCAGATTATC 2180
642 .....LeuAspSerSerCysLysGlnSerSerVal 651
      |||||:::|:::|:::|
2181 AACAGCAACAGTCCTTGGCTCTGACTCAAGCTGTATGCT..... 2220
652  SerValThrSerProSerGlyValSerSerSerThrSerGlyValSe 668
      :::|:::|:::|:::|:::|:::|
2221 ACAGAGAACACCTACATCTAGACAGCAACAAATCACCACAAAGCAAG 2270
668  rSerThrSerAsnMetHisGlySerLeuLeuGlnGluLysHisArgIle 684
      |||||:::|:::|:::|
2271 TAGTAGGGGACCAACAGCAACAGCTGTGTGAGAAAGTAAAGCAACAGCA 2320
685  LeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAlaLysI 701
      :::::|:::|:::|:::|
2321 GTTAATGGCATGTTGAATGGAATCTAACCAAAATCGTCCTTCA 2370
701  eThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysGlyA 718
      |||:::|:::|:::|
2371 TTGTCCACAG.....CAAGACCCACAAACATATATATGCTTTACAG 2411
718  spGlyAsnValValLysGlnGlnGlnLeuSerProLysLysGlnLysn 734
      ||:::|:::|:::|
2412 ACTTACCTGGGATCAGTCAAGAG.....TTCCCTTACAAATCTGAAGATG 2455
735  AsnAlaLeu.....LeuArgTyrLeuLeuAspArgAspAspProSerAs 749
      :::::|:::|:::|
2456 GATTCTATGCGCTTATACACAGAACTTATTTCCGTATATACGCT.... 2500
749  pAlaLeuSerLysGlnLeuGlnProGlnValGluGlyValAspAsnLysm 766
      :::|:::|:::|
2501 .....GTATTACCACAA.....C 2513
766  eIleGlnGlyThrSerSerThrIleProSerSerSerGlnGlnLysAsp 782
      |||||:::|:::|
2514 ATTCCAAATATACAGAGCTGACTTACCTATGGGAGTTTGTGAACCTCC 2563
783  ProlLysIleLysThrGlnThrSerGlnGlnGlySerGlyAspLeuAspAs 799
      |||:::|:::|
2564 CCAATACCCCACTACTCTAGTTTGAAG..... 2590
799  nLeuAspAlaIleLeuGlnLysLeuThrSerSerAspPheTyrAsn...A 815
      |||||:::|
2591 .....GATTTTGTCACTTGT 2606

```

```

815  snSerIleSerAsnGlySerHisLeuGlyThrLysGlnGlnValPhe 831
      :::::|:::|:::|
2607 TACACTTCTGTAACCAACCAAGCAT..... 2632
832  GlnGlyThrAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerI 848
      |||||:::|:::|
2633 .....GGATTAAATCCACAGTCAACCATTAATTAATCACTCC 2664
848  eArgProProTyrAsnArgAlaValSerLeuAspSerProValSerValG 865
      :::|:::|:::|
2665 TCAGACATGTTATGCTGGGCGCGTGTGATGATCACTGACACAGCAAC 2714
865  LysSerSerProProValLysAsnIleSerAlaPheProMetLeuProl 881
      :::::|:::|:::|
2715 CTCAGCACACCCACAGTGGGTGAGTGCAGTACATCCAGTACGCGCC 2764
882  GlnProMetLeuGlyLysAsnProArgMetMetAspSerGlnGlnLysn 898
      |||:::|:::|
2765 CAACAGCATTTTAAACAGATTCCAGATGAGAGTTTAATGAAACATA 2814
898  rGlySerSerMetGlyGlyProAsnArgAsnValThrValThrGlnThr 915
      |:::|:::|
2815 TCCAGCTGAATTA.....ATAACATAAATMACCTCAGACTA 2852
915  rSerSerGlyAspTyrGlyLeuProAsnSerLysAlaGlyArgMetGlu 931
      :::::|
2853 CCACA.....CATCTTCAG 2866
932  ProMetAsnSerAsnSerMetGlyArgProGlyLysPtyrAsnThrSe 948
      |||||:::|
2867 CCACTTATATATCCGTGACGAAGACGACCTTTTCGTGATTTGACATCCAG 2916
948  r 948
2917 T 2917

```

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-366-051B-3

```

seq_documentation_block:
; Sequence 3, Application US/08366051B
; Patent No. 5650283
; GENERAL INFORMATION:
; APPLICANT: Bradwick, Christopher A.
; APPLICANT: Dolwick, Kristin M.
; APPLICANT: Carver, Lucy A.
; TITLE OF INVENTION: Ah Receptor cDNAs and Genetically
; TITLE OF INVENTION: Engineered cells for Detecting Agonists to the Ah
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESS: Tilton, Fallon, Lungmus & Chestnut
; STREET: 100 South Wacker Drive, Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/366, 051B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Tilton, Timothy L.
; REGISTRATION NUMBER: 16,926
; REFERENCE/DOCKET NUMBER: NU-9207-CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)-456-8000
; TELEFAX: (312)-456-7776
; INFORMATION FOR SEQ ID NO: 3:

```

```

SEQUENCE CHARACTERISTICS:
LENGTH: 5261 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 383..2927
US-08-366-051B-3

```

```

alignment_scores:
Quality: 252.00 Length: 1051
Ratio: 0.545 Gaps: 43
Percent Similarity: 43.958 Percent Identity: 19.696

```

```

alignment_block:
US-09-041-994-2 x US-08-366-051B-3 ..

```

```

Align seg 1/1 to: US-08-366-051B-3 from: 1 to: 5261

```

```

15 SerArglyArgLysLeuProCysASP.....ThrProgl 26
|||||
416 AGTCGCAAGCGCGGAGCCGCTGCAGAAAACAGTAACCAATCCAGC 465
26 yGlnGlyLeuThrCysSerGlyLysArgArgArgGlnGlnSerL 43
|||||
466 TGAAGGAATCAAGTCMAATCTTCMAAGCGCATAGACGACTTAATA 515
43 yGlyTrlLeuGlnLeuAlaGlnLeuLieserAlaAsnLeuSerAsp 59
|||||
516 CAGAGTTGACCGCTTGGCTAGCTGCTGCTTCCCA...CAGATGTT 562
60 AspAsnPheAsnValLysProAspLysCysAlaLeuLeuLysGlnThra 76
|||||
563 ATTAAAT.....AAGTTGGACAAACTTTCAGTCTTAGCGTCGCT 603
76 lArgGlnLeuArg.....GlnlLeysGlnGlnG 86
|||||
604 CAGTTACCTAGAGACCCAGAGCTTCTTGATGTTGCATTAAATCCTCCC 653
86 LysLysThrLieserAsnAspAspValGlnLysAlaAspValSerSer 102
|||||
654 CTACTGAAAGAAAGAGGAGCCAGGATACTGTAGAGCAGCAAAATTTACA 703
103 ThrGlnGlnGlyValLLeasPlysAspSerLeuGlyProLeuLeuGln 119
|||||
704 GAAGGCGTGAAGCTTACAGAA.....GGAGATTTCTTATTATACA 741
119 nAlaLeuAspGlyPheLeuPheValValAsnArgGlnAlaAsnLeuValP 136
|||||
742 GGCCTCGAATGGCTTGTATGTTGTCACAGATGCTTGGCTTTT 791
136 heValSerGlnAsnValThrGlnTyTrLeuGlnTyTrLysGlnGlnLysPleu 152
|||||
792 ATGCTTCTTCTACTATACAAGATATCTAGGGTTTCAGCAGTCTGATGTC 841
153 ValAsnThrSerValTyTrAsnLLeuLiesGlnGlnLysAspArgLysAspH 169
|||||
842 ATACATCAGAGTGATATGATCACTTATCCATACCAGAGCAGAGCTGAAT 891
169 eLeuLysAsnLeuProLysSerThrValAsnGlnLysLertPrpThrAsnG 186
|
892 T..... 892
186 lUrProGlnArgLysSerHisThrPheAsnGlyArgMetLeuMetLys 202
|||||
893 .....CAGGCTCAGCTACACTGGGCAATTAAAT..... 919
203 ThrProHisAspTrlLeuGlnLysPleuAsnAlaSerProGlnMetArgL 219
|||||
920 ...CCTTTCAGTGTACAGAGTCTGAGCAAGGAATTAAGAAAGCCACCTGG 966

```

```

219 nArgTyTrGlnThrMetGlnCysPheAlaLeuSerGlnProArgAlaMetL 236
|||||
967 TCTCCCCCAGACAGTAGTCTGTTAT.....AACCAGACAGATTC 1007
236 eGlnGlnGlnGlnLysPleu...GlnSerCysMetLLeCysValAlaArg 251
|||||
1008 CTCAGAAAACCTCTCTTAAATGAGAGAGTGCTTCATATGTCGTAAAG 1057
252 ArgLethrThrGlyLysArgThrPhe..... 260
|||||
1058 TGTCTGCTGGATAATTCACTGTGTTTCTGGCAATGAATTTCCAAAGGAA 1107
260 ..... 260
1108 GTTAAAGTATCTTCATGCGACAGAAAAAGAAAGATGATGATCAATAC 1157
261 .....ProSerAsnProGlu 265
|||||
1158 TTCACCTCAGTTGGCTTTGTTGGATAGCTACCTCAGCCACCA 1207
266 Ser.....PheLethrArgHis.. 271
|||||
1208 TCCATACTTGAATCCGACCAAAATTTTATCTTAGACCAACCAACA 1257
272 .....AspLeuSerGlyLysValValAsnI 280
|||||
1258 ACTAGACTTCACACCTATTGTTGTCATCCCAAGAAAGAAATTTGTTAG 1307
280 lAspPThrAsnLeuArgSerSerMetLArgProGlyPheGlnAspL 296
|||||
1308 GATATATGAGACAGAGCTGTGCACGAGGCTCAGGTTATCAGTTTAT 1357
297 .....lLeuArg..... 298
1358 CATGACGCTGATATGCTTATTTGCGGAGATGCCATATCCGAATGATTAA 1407
299 .....ArgCyslLeuArgPhePheSerLeuAsnAspG 310
|||||
1408 GACTGGAGAAAGTGCATGATGTTTCCGCTCTTCAAAAAACAAC 1457
310 lGlnSerTrpSerGlnLysArgHisTyTrGlnGlnAlaTyTrLeuAsnGly 326
|||||
1458 GATGACCTTGGCTCCAGCT...AATGCAGCCTGCTTTATAAAAATGGA 1504
327 HisAlaGlnThrProValTyTrArgPheSerLeuAlaAspGlyThrLLeva 343
|||||
1505 AGA.....CCAGATTATATCATTTGT 1524
343 lThrAlaGlnThrLysSerLysLeuPheArgAsnProValThrAsnAspA 360
|||||
1525 AACT.....CAGAGACCACTAACAGATGAGG 1550
360 rGHisGlyPheValSerThrHisPheLeuGlnArgGlnAsnGlyTyTr 376
|||||
1551 AA.....GGAACAGCATTTA...CGAAACGAATATCGAAG 1585
377 ArgProAsnProAsnProValGlnGlnGlnLysLLeaArgProProMetAlaG 393
|||||
1586 TTGCTTTTATGTTTAAACCACTGAGAAAGCTGTGTATGAGCAACCA 1635
393 yCysAsnSerSerValGlyGlyMetSerMetSerProAsnGlnGlnLys 410
|||||
1636 CCGTTTTCCTGCCATA.....ATGATGCC.....TTAC 1664
410 lMetProSerSerArgAlaTyTrGlyLeuAlaAspProSerThrThrGly 426
|||||
1665 CACTAAGGACTAAAAATGSCACTAGTGAAGAAAGATCTGCTCCACATCC 1714
427 GlnMetSerGlyAlaArgTyTrGlyLysSerSerAsnLLeaSer..... 441
|||||
1715 ACTTAAGCAAGAGACTCTCTCAATCTTAATCTCCCTGGGCTGCATGAT 1764

```



```

442 .....LeuThrProGlyProGlyMetGlnSerP 451
1765 GCACAAAGATGAGCTGTTATCTCTATCTGCTTCAAGTACTTCAAGTA 1814
451 roSerSerTyrglnAsnAsnAsnTyrGlyLeuAsnMetSerSerPro 467
1815 CTGCACCTTTGAAACAACTTTTTCACGAACTATGATGATGATGAGA 1864
468 HisGlySerProGlyLeuAlaProAsnGlnGlnAsnIleMetIleSerP 484
1865 AATTGGCAAGATATACTGCACCG..... 1888
484 oArgAsnArgGlySerProLysIleAlaSerHisGlnPheSerProValA 501
1889 .....ATGGGAAATGATACTATCTCGAAACATGAG..... 1918
501 IacGlyAlaHisSerProMetAlaSerSerGlyAsnThrGlyAsnHisSer 517
1919 ..CAAAATTGACACCTCAGAGATGTGAACATCATTTGCTGGAGGTCCACA 1966
518 .....PheSerSerSerSerLeuSerAlaLeuGlnAlaIleSerGluG 532
1967 GGGCTCTTCAGATAGTAAACACGATGACTTGTACAGCATAAATGAAAAA 2016
532 yValGlyThrSerLeuLeuSerThrIleSerSerProGlyProLysLeuA 549
2017 CTTAGGCAATTGATTTT..... 2032
549 sPaAsnSerProAsnMetAsnIleThrGlnProSerLysValSerAsnGln 565
2033 .....GAAGACATCAGACACATGCAGATGAAATTTTTCAGAAAT 2074
566 AspSerLysSerProLeuGlyPheTyrCysAspGlnAsnProValGluSe 582
2075 GATTTTCTGTGTGAGCTTGACTTC..... 2098
582 rSerMetCysGlnSerAsnSerArgasp...HisLeuSerAspLysGluS 598
2099 .....AGAGACATTTGACTTAAACGAT..... 2119
598 eArgLysGlnSerSerValGluGlyAlaGlnAsnGlnArgGlyProLeuGlu 614
2119 ..... 2119
615 SerLysGlnHisLysLysLeuLeuGlnLeuThrCysSerSerAspAs 631
2120 .....GAAATCTCTGACGTATGTCCAAAGAT... 2143
631 parGlyHisSerSerLeuThrAsnSerPro..... 641
2144 .....TCTTTAAGTAACTCTCCCTTCATACCTTCAGATTATC 2180
642 .....LeuAspSerSerCysLysGlnSerSerVal 651
2181 AACGCAACAGTCTTGGCTCTGAACTCACAGCTTATGCT..... 2220
652 SerValThrSerProSerGlyValSerSerSerThrSerGlyGlyAlaSe 668
2221 ACAGGAACACCTACATCTAGAACAGCAACACATCACCACCAAGCAAG 2270
668 rSerThrSerAsnMetHisGlySerLeuLeuGlnGluLysHisArgIle 684
2271 TAGTAGTGGAGCCACAGCAACAGCTGTGTGAGAAATGATGAAGACATGC 2320
685 LeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAlaLysII 701
2321 GTTAATGGCATGTTGAAAAATGGAACCTTAACAACAAATCGCTTCACA 2370
701 eThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysGlyA 718
2371 TTGTCCACAG.....CAMGACCCACAAACAATATATGCTTTTACAG 2411
718 spGlyAsnValValLysGlnGlnLeuSerProLysLysLysGlnAsn 734

```

```

11 ..... 1111111111
2412 ACTTACATGGGATGATCCAAAG.....TTCCCTCAAAATCTGAATG 2455
735 AsnAlaLeu.....LeuArgTyrLeuLeuAspArgAspAspProSerAs 749
2456 GATTCATGCCCTTATACACAGAACTTATTTCTGTAATCAGCCT..... 2500
749 palLeuSerLysGlnLeuGlnProGlnValGluGlyValAlaAspAsnLysM 766
2501 .....GATTTACCAACA.....C 2513
766 eSerGlnCysThrSerSerThrIleProSerSerSerGlnGluLysasp 782
2514 ATTCCAATGTACAGACCTGACATCCCTATGGGAGATTTTGAACCTCC 2563
783 ProLysIleLysThrGlnThrSerGluGluGlySerGlyAspLeuAspAs 799
2564 CCATACCCCACTACTCTAGTTTAGAA..... 2590
799 nLeuAspAlaIleLeuGlnLysPheLeuThrSerSerAspPheTyrAsn...A 815
2591 .....GATTTGTCACTTGTT 2606
815 snSerIleSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValPhe 831
2607 TACACTTCTCGAAACCAAAAGCAT..... 2632
832 GlnGlyThrAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerII 848
2633 .....GGATTAATTCACACAGTCCATCAATATATCCTCC 2664
848 eArgProProTyrAsnArgAlaValSerLeuAspSerProValSerValG 865
2665 TCAACATGTTATGCTGGGCGGTGATGTATCACTGCCAGCCAGCAAC 2714
865 LysSerSerProProValLysAsnIleSerAlaPheProMetLeuProLys 881
2715 CTCAGCACACCCACGTGGGTGACATGCATCAACATCCAGTACTGCCAGCC 2764
882 GluProMetLeuGlyGlyAsnProArgMetMetAspSerGlnGluAsnTy 898
2765 CAACAGCATTTTAAACAAGTTTCAGATGAGACTTTTAATATGAACATA 2814
898 rGlySerSerMetGlyGlyProAsnArgAsnValThrValThrGlnThrP 915
2815 TCCAGCTGAATTA.....AATAACATTAATATACACTCAGACTA 2852
915 roSerSerGlyAspTyrPglyLeuProAsnSerLysAlaGlyArgMetGlu 931
2853 CCACA.....CATCTTCAG 2866
932 ProMetAsnSerAsnSerMetGlyArgProGlyGlyAspTyrAsnThrSe 948
2867 CCATTTATATCATCGTCAAGAACCAAGACCTTTTCTGATTTGACATCCAG 2916
948 r 948
2917 T 2917

```

51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67
 |||||
 236 CTGATATCTGCCAATCTTAGTGATATTGCAATTTCAATGTCAAAACAGA 285
 |||||
 67 pLysGlyAlaIleLeuLysGluThrValArgIleArgIleLysG 84
 |||||
 286 TAAATGTGGGATTTTAAAGGAACAGTAAGACATACGTCAAATTAAGA 335
 |||||
 84 IuGInGlySerThrIleSerAsnAspAspValGInLysAlaAspVal 100
 |||||
 336 AGCAAGGAAAACTATTCCATGATGATGATGTTCAAAAAGCCGATGTA 385
 |||||
 101 SerSerThrGlyGInGlyValIleAspLysAspSerLeuGlyProLeu 117
 |||||
 386 TCTTGTACAGGGCAGGAGATTATTGATTAAGACTCTTAGACCCCTTTT 435
 |||||
 117 uLeuGInAlaLeuAspGlyPheLeuPheValValAsnArgIuAlaAsnI 134
 |||||
 436 ACTTACGGCATTGGATGGTTCTATTGTGGTGATGACGAGGCAACAA 485
 |||||
 134 LevalPheValSerGluAsnValThrGInLysLeuGInLysGInGlu 150
 |||||
 486 TTGTATTGTATACAAAAATGCACACAATACCTGCAATATATAGCAAG 535
 |||||
 151 AspLeuValAsnThrSerValLysAsnIleLeuHisGluLysAspArg 167
 |||||
 536 GACCTGGTTAACCAAGTGTTTACATATCTTACATGAAGAAGACAGAA 585
 |||||
 167 sAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrp 184
 |||||
 586 GGATTTTCTTAAGAAATTACCAAAATCTACAGTTAATGAGATTCTCTG 635
 |||||
 184 hrAsnGluProGInArgGInLysSerHisThrPheAsnGlyArgMetLeu 200
 |||||
 636 CAAATAGAGCCCAAGACAAAAACCATACATTAATGCGCGTATGTTG 685
 |||||
 201 MetLysThrProHisAspIleLeuGInLysPheAsnAlaSerProGlu 217
 |||||
 686 ATGAATAACACCATGATATTCTGGAAGACATTAACGCCAGTCCGAAT 735
 |||||
 217 LArgGInArgGlyGluThrMetGInCysPheAlaLeuSerGInProArg 234
 |||||
 736 GCGCCAGAGATATGAACAATGCACTGCTTCCCTGTCTCAGCCACGAG 785
 |||||
 234 IametMetGluGluGlyLysAspLeuGInSerCysMetIleCysValAla 250
 |||||
 786 CTATGATGAGGAAGGGAAGATTGCAATCTTGTATGATCTGTGGCA 835
 |||||
 251 ArgArgIleThrThrGlyGluArgThrPheProSerAsnProGluSerP 267
 |||||
 836 CGCCGCAATTACTACAGGAAGAAACATTTCCATCAAAACCCGGAAGCT 885
 |||||
 267 eIleThrArgHisAspLeuSerGlyLysValValAsnIleAspThrAsn 284
 |||||
 886 TATTCCAGACATGATCTTTCCAGAAAGTTTCTCAATATAGTACAAATT 935
 |||||
 284 erLeuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300
 |||||
 936 CACTGAGATCTCCATGAGGCTGCTTGAAGATATATATCCGAAGGTGT 985
 |||||
 301 IleGInArgPhePheSerLeuAsnAspGlyGInSerTrpSerGInLysArg 317
 |||||
 986 ATTCAAGATTTTAACTTAATATGATGGCGCATGATGTCACAAACG 1035
 |||||
 317 gHisTrpGInGluAlaLysLeuAsnGlyHisAlaGluThrProValLys 334
 |||||
 1036 TCACATACAAAGAGCTTATCTTAATGCGCATACAAACCCAGATATATC 1085
 |||||
 334 rgPheSerLeuAlaAspGlyThrIleValThrAlaGInThrLysSerLys 350
 |||||
 1086 GATTCGTGTTGGCTGATGAACTATAGTACGACAGACAAAGCAAA 1135

351 LeuPheArgAsnProValThrAsnAspArgHisGlyPheValSerThrHis 367
 |||||
 1136 CTCTCCGGAATCTGTAAACAATGATGACATGCTTGTGTCGTAACCCA 1185
 |||||
 367 sPheLeuGInArgGluGInAsnGlyThrArgProAsnProAsnProValG 384
 |||||
 1186 CTTCCTTCAGAGAGAACAAATGATATAGACCAAAACCAATCTCTGTG 1235
 |||||
 384 IyGInGlyIleArgProPheMetAlaGlyCysAsnSerSerValGlyGly 400
 |||||
 1236 GACAAGGGATTAGACCACTATGCTGATGCAACAGTTGCTAGCGGCG 1285
 |||||
 401 MetSerMetSerProAsnGInGlyLeuGInMetProSerSerArgAlaTy 417
 |||||
 1286 ATGAGTATGTCCGCAACCAAGCTTACAGATGCCAGACAGAGGCCCTA 1335
 |||||
 417 rGlyLeuAlaAspProSerThrThrGlyGInMetSerGlyAlaArgTyG 434
 |||||
 1336 TGGCTTGGCAGACCCTAGCACCAAGGCGAGATGATGAGCTTAGGTATG 1385
 |||||
 434 IyGlySerSerAsnIleAlaSerLeuThrProGlyProGlyMetGInSer 450
 |||||
 1386 GGGGTTCCAGTACATAGCTTATGACCCCTGGGCGAGGATGCAATCA 1435
 |||||
 451 ProSerSerTyrgInAsnAsnAsnTyrgIleAsnMetSerSerProTr 467
 |||||
 1436 CCATCTTCTTACAGAACCAACATATGGGCTCAACATGATAGAGCCCC 1485
 |||||
 467 oHisGlySerProGlyLeuAlaProAsnGInGInAsnIleMetLieserP 484
 |||||
 1486 ACATGGAGTCTGTGTTGCTTGGCCCAACACGAGAAATCATGATTTCTC 1535
 |||||
 484 roArgAsnArgGlySerProLysIleAlaSerHisGInPheSerProVal 500
 |||||
 1536 CTCTTAATCTGTGGAGTCCAAAGATAGCCTCACAATGATTTCTCTGTT 1585
 |||||
 501 AlaGlyValHisSerProMetAlaSerSerGlyAsnThrGlyAsnHisSe 517
 |||||
 1586 GCGGTGTGCACTCTCCATGGCATCTTCTGGCAATACTGGGAACCAAG 1635
 |||||
 517 rPheSerSerSerSerLeuSerAlaLeuGInAlaIleSerGluGlyValG 534
 |||||
 1636 CTTTTCACAGACCTCTCTCAGTCCCTGCAAGCCATCTCAGTAGGTGG 1685
 |||||
 534 IyThrSerLeuLeuSerThrLeuSerSerProGlyProLysLeuAspAsn 550
 |||||
 1686 GGACTTCCCTTTATCTACTCTGTCATGCACAGGCCCAAAATTGATTAAC 1735
 |||||
 551 SerProAsnMetAsnIleThrGInProSerLysValSerAsnGInAspSe 567
 |||||
 1736 TCTCCCAATATGAATATTACCCCAAGTAAGTAAGTAACATCAGGATTC 1785
 |||||
 567 rLysSerProLeuGlyPheTyrcysAspGInAsnProValGluSerSerm 584
 |||||
 1786 CAAGAGTCTCTGGGCTTTTATTTGCGACCAAAATCAGTGGAGAGTTCAA 1835
 |||||
 584 eCysGInSerAsnSerArgAspHisLeuSerAspLysGluSerLysGlu 600
 |||||
 1836 TGTGTCACTCAAAATAGCAAGATCACCTCAGACAAAGAAAGTAAGAG 1885
 |||||
 601 SerSerValGluGlyAlaGluAsnGInArgGlyProLeuGluSerLysG 617
 |||||
 1886 AGCAGTGTGAGGGGGCAGAGAAATCAAAAGGGGTCTTTGGAAAGCAAAAG 1935
 |||||
 617 yHisLysLysLeuLeuGInLeuLeuThrCysSerSerAspAspArgGlyH 634
 |||||
 1936 TCATAAAAAATTACTGCAATTACTTACTGTTCTTCTATGACCGGGGCTC 1985
 |||||
 634 IAspSerLeuThrAsnSerProLeuAspSerSerCysAlaGluSerSer 650
 |||||
 1986 ATTCTCTCTTGACCAACTCCGCCCTAGATTCAAAGTTTAAAGAAATCTTCT 2035
 |||||
 651 ValSerValThrSerProSerGlyValSerSerSerThrSerGlyGlyVal 667

```

|||||
2036 GTTAGTGTACACAGCCCCCTCTGGAGTCTCTCTCTACATCTGGAGAGCT 2085
|||||
667 lSerSerThrSerAsnMetHisGlySerLeuLeuGlnGlnIuLysHisArgI 684
|||||
2086 ATCTCTACATCCAAATATGATGGGTGCTACTGTCTACAGAGAACAGACCGGA 2135
|||||
684 lLeuHisLysLeuLeuGlnAsnGlnLysSerProAlaGluValAlaLys 700
|||||
2136 TTTTGGCAAGTTGCTCGCAATGGGAATTCACACAGCTGAGGTAGCCAG 2185
|||||
701 lIleThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysGlu 717
|||||
2186 ATTAGTGTACAGAGCCCTGGAGAAAGACACAGAGTATACCTTCTGTGG 2235
|||||
717 yAspGlyAsnValAlaLysGlnGlnLeuSerProLysLysGlu 734
|||||
2236 GGACGGAAATGTTGTCAAGCAGAGCAGCTAAGTCTAAGAAAGAGAGA 2285
|||||
734 snAsnAlaLeuLeuArgTyrLeuLeuAspArgAspProSerAspAla 750
|||||
2286 ATAAATGCATCTTAGATACCTGCTGACAGGAGATGATCCTAGTGATGCA 2335
|||||
751 LeuSerLysGlnLeuGlnProGlnValGluGlyValAspAsnLysMetSe 767
|||||
2336 CTCTCTAAAGACTACAGCCCAAGTGGAGAGTGGACAATAAATGAG 2385
|||||
767 rGlnCysThrSerSerThrIleProSerSerSerGlnIuLysAspProL 784
|||||
2386 TCAGTGCACCGCTCCACCATTTCTAGCTCAAGTCAAGAGAAAGACCCTA 2435
|||||
784 ySileLysThrGlnThrSerGlnGlnLysSerGlyAspLeuAspAsnLeu 800
|||||
2436 AAATTAAGACAGACAGAGTGAAGAGGATCTGGAGACTGGATGAATCTA 2485
|||||
801 AspAlaIleLeuGlyAspLeuThrSerSerAspPheTyrAsnAsnSerI 817
|||||
2486 GATGCTATTTCTGTGTGATCTGACTAGTTCCTTACATTAACATATTCAT 2535
|||||
817 eSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValPheGlnGlyT 834
|||||
2536 ATCTCTAAATGTAGTCACTGTGGGAGCTAAGCAACAGGTTTCAAGAA 2585
|||||
834 hrAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleArgPro 850
|||||
2586 CTAAATCTCTGGTTGAAAGTTGACAGTCTGACAGTCAATTCGTCCT 2635
|||||
851 ProTyrAsnArgAlaValSerLeuAspSerProValSerValGlySerSe 867
|||||
2636 CCATATAACCGAGCAGTGTCTCTGATRAGCCCTGTTTCTGTGGCTCAG 2685
|||||
867 rProProValLysAsnIleSerAlaPheProMetLeuProLysGlnProm 884
|||||
2686 TCCTCCAGTAAATAATATCACTGCTTCCCATGTACCAAGAACACCA 2735
|||||
884 eLeuGlnGlyAsnProArgMetMetAspSerGlnLysAsnTyrGlySer 900
|||||
2736 TGTGGGTGGGAATCCAAAGATGATGATAGTACGGAATAATTAAGGCTCA 2785
|||||
901 SerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProSerSe 917
|||||
2786 AGTATGGGTGGGCCAAACCGAAATGTGACTGTGACTCAGACTCCTTCTC 2835
|||||
917 rGlyAspTyrGlyLeuProAsnSerLysAlaGlyArgMetGluProMetA 934
|||||
2836 AGGAGACTGGGGCTTACCAGAACTCAAGGCGGAGAGATGAGAACCTATGA 2885
|||||
934 snSerAsnSerMetGlyArgProGlyArgTyrAsnThrSerLeuPro 950
|||||
2886 ATTCAAACTCCATGGAGACAGCAGAGAGATTAATTAATCTTTTACCC 2935
|||||
951 ArgProAlaLeuGlyGlySerIleProThrLeuProLeuArgSerAsnSe 967
|||||
2936 AGACCTGCACCTGGGTGGCTATTTCCACATTTGGCTCTTGGCTAATAG 2985
|||||
967 rIleProGlyValaArgProValLeuGlnGlnGlnGlnMetLeuGlnm 984
|||||
2986 CATACACAGGTGCGAGACAGATTTGCAACACAGACAGACAGATGCTTCAAA 3035
|||||
984 eTargProGlyGluIleProMetGlyMetGlyAlaAsnProTyrGlyGln 1000
|||||
3036 TGAGGCGCTGGTGAATCCCGCATGGGAATGGGGGCTTAATCCCTATGGCCAA 3085
|||||
1001 AlaAlaIleSerAsnGlnLeuGlySerTyrProAspGlyMetLeuSerMe 1017
|||||
3086 GCACAGCATTTAAACCAACTGGGTTCCTGGCCCATGAGCATTTGTCCAT 3135
|||||
1017 fGluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerI 1034
|||||
3136 GGAACAAGTTTCTCATGCGACTCAAAATAGGCTCTTCTTAGAATTTCC 3185
|||||
1034 euAspAspLeuValGlyProProSerAsnLeuGlnGlyGlnSerAspGlu 1050
|||||
3186 TGATATGATCTTTGTGGCCACCTTCCAACTGGAAGGCCAGAGTGAGGAA 3235
|||||
1051 ArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaTh 1067
|||||
3236 AGACCATTTATGGACCACTGACACACTTCTCTCAGCAACACAGATGGCAG 3285
|||||
1067 rGlyLeuGlnGlnIleAspArgAlaLeuGlyIleProGluLeuValAsnG 1084
|||||
3286 AGGCGCTGGAGAAATTTGACAGAGCTTGGCGATCTCGAATCTGTCAATC 3335
|||||
1084 lngGlnAlaLeuGlnProLysGlnAspAlaPheGlnGlnGlnAla 1100
|||||
3336 AGGACAGCGCATTTAGAGCCCAACAGAGATGCTTTCAGAGGCCAAGAAACA 3385
|||||
1101 AlaValMetMetAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAl 1117
|||||
3386 GCATTAATGATGATGATCAGAAAGCAGGATTAATGAGACAGATACCAGC 3435
|||||
1117 gGlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlyGlnSerProS 1134
|||||
3436 ACAGGGGCTTCAATGCAAGGAGGCTTTCATCTCAGGACCAATCACCAT 3485
|||||
1134 erPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlyAsnPheProLeu 1150
|||||
3486 CTTTAACTCTATGATGATCAGATGAACACAGAGGCAATTTTCTCTCTC 3535
|||||
1151 GlnGlyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPr 1167
|||||
3536 CAAGGAATGCACCCAGCAGCCACATCATGTGAGACCCCGAGCAAAACACCC 3585
|||||
1167 oLysGlnLeuArgMetGlnLeuGlnGlnArgLeuGlnGlyGlnPheL 1184
|||||
3586 CAACCAACTTGAATGACAGCTTACAGCAGAGGCTCAGAGGCCAGACGATTT 3635
|||||
1184 euAsnGlnSerArgGlnAlaLeuGlnLysMetGluAsnProThrAla 1200
|||||
3636 TGAATCAGAGCGCAGAGCCTTAATTAATGAATAAGAAACCCCTACAGCT 3685
|||||
1201 GlyGlyAlaAlaValMetArgProMetMetGlnProGlnGlnGlyPhe 1217
|||||
3686 GGTGGTGTGGGCGATGAGGCTATGATGACAGCCCAAGAGGGTTTCTT 3735
|||||
1217 uAsnAlaGlnMetValAlaGlnArgSerArgGluLeuLeuSerHisHisP 1234
|||||
3736 TTAATGCTCAAAATGGTGGCCCAACGACAGAGAGACTGCTAATGATCACT 3785
|||||
1234 hearGlnGlnArgValAlaMetMetMetGlnGlnGlnGlnGlnGln 1250
|||||
3786 TCCGACACAGAGGGGTGGTATGATGATGACAGCAGCAGCAACAGCAGAG 3835
|||||
1251 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1267
|||||
3836 CAGCAGCAGCAGCAGCAGCAACACAGCAACAGCAGCAGCAACAGCAGCA 3885

```

```
1267 nThrglnAlaPheSerProProAsnValThrAlaSerProSerMet 1284
      |||
3886 AACCCAGGCTTCAGCCACCTCTATGTGACTGCTCCAGCATGG 3935
1284 sRGlyLeuLeuAlaGlyProThrMetProGlnAlaProProGlnGlnPhe 1300
      |||
3936 ATGGGCTTTGGGAGGCCACACATGCCACAGCTCTCCGCAACAGTTT 3985
1301 ProTyrGlnProAsnTyrGlyMetGlyGlnGlnProAspProAlaPheG1 1317
      |||
3986 CCATATTCACCAATATATGGAATGGGCAACAACGATCCAGCCTTTGG 4035
1317 yArgValSerSerProProAsnAlaMetCetSerSerArgMetGlyPro 1334
      |||
4036 TCGAGGTCTAGTCTCCCAATGCATGATGTCGTCAAGAAATGGGTCCT 4085
1334 eRGlnAsnProMetMetGlnHisProGlnAlaAlaSerLLeTyrGlnSer 1350
      |||
4086 CCCAGATCCCATGATGACACACCCGACGGTCATCCATCTATCAGTCC 4135
1351 SerGluMetLysGlyTyrProSerGlyAsnLeuAlaArgAsnSerSerPh 1367
      |||
4136 TCAGAAATTAAGGCTGGCCATCAGAAATTTGGCCAGGACAGCTCTT 4185
1367 eSerGlnGlnGlnPheAlaHisGlnGlyAsnProAlaValTyrSerMetV 1384
      |||
4186 TTCACAGCAGCAGTTTCCACACAGGGAATCTCGCAGTGTATGATGG 4235
1384 aHisMetAsnGlySerSerGlyHisMetGlyGlnMetAsnMetAsnPro 1400
      |||
4236 TCACACTGATGAGCAGACAGTGTACATGAGACAGATGAATGAACCCC 4285
1401 MetProMetSerGlyMetProMetGlyProAspGlnLysTyrCys 1415
      |||
4286 ATGCCCATGTCTGGCATGGCTATGGTCTGTGATCAAAATACTGCG 4330

seq_name: /cgn2_6/p/tdodata/1/pna/us095a_COMB.seq:us-09-513-066-1

seq_documentation_block:
: Sequence 1, Application US/09513066
: GENERAL INFORMATION:
: APPLICANT: Chen, J. Don
: APPLICANT: Leo, Christopher
: APPLICANT: Li, Hai
: TITLE OF INVENTION: NOVEL USES FOR THE RAC3 TRANSCRIPTIONAL COACTIVATOR OF
: FILE REFERENCE: UMG-026CP
: CURRENT APPLICATION NUMBER: US/09/513,066
: PRIOR APPLICATION NUMBER: USSN 09/041,994
: PRIOR FILING DATE: 1998-03-13
: PRIOR APPLICATION NUMBER: USSN 60/073,674
: PRIOR FILING DATE: 1998-02-04
: NUMBER OF SEQ ID NOS: 37
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 4496
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (86)..(4330)
US-09-513-066-1

alignment_scores:
      Quality: 7375.00      Length: 1415
      Ratio: 5.216      Gaps: 0
      Percent Similarity: 99.929      Percent Identity: 99.929

alignment_block:
US-09-041-994-2 x US-09-513-066-1
```

```
Align seg 1/1 to: US-09-513-066-1 from: 1 to: 4496
1 MetSerGlyLeuGlyGlnAsnLeuAspProLeuAlaSerAspSerArgLy 17
      |||
86 ATGAGTGGATTAGAGCAAAACCTTGATCCATCGCCAGATTCACAGAAA 135
17 sArgLysLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGlyG 34
      |||
136 ACGCAAAATGGCATGTGATATCTCCAGGACAAAGTCTTACCTCAGAGTGG 185
34 LuLysArgArgArgGlnGlnGlnLysTyrTrLLeGlnLLeuAlaGlu 50
      |||
186 AAAAAGGAGAGGGGAGCAGAAAGTAATTAATTAACAAATTTGGCTGAG 235
51 LeuLLeSerAlaAsnLeuSerAspLLeAspAsnPhAsnValLysProAs 67
      |||
236 CTGATATCTGCCCAATCTTATGTATATGCAATTTCAATGTCAAAACACA 285
67 pLysCysAlaLLeLeuLysGluThrValArgGlnLLeArgGlnLLeLysG 84
      |||
286 TAAATGTGGGATTTTAAAGCAACAGTAAGACAGATACGTCAATTAAGA 335
84 LuGlnGlyLysThrLLeSerAsnAspAspAspValGlnLysAlaAspVal 100
      |||
336 ACGCAAGGAAAACTATTTCCAATGATGATGATTTCAAAACCCGATGTA 385
101 SerSerThrGlyGlnGlyValLLeAspLysAspSerLeuGlyProLeu 117
      |||
386 TCTTTCTACAGGGCAGGAGCTTATTCATTAAGACTCTTGAGACCGCTTT 435
117 uLeuGlnAlaLeuAspGlyPheLeuPheValValAsnArgGlnAlaAsnI 134
      |||
436 ACTTGAGGATTTGATGATGGTTCTTATTTGGTGATTCAGAGGCAACA 485
134 LeValPheValSerGluAsnValThrGlnTyrLeuGlnTyrLysGlnGlu 150
      |||
486 TTGTATTTGTATCAGAAATGTCACACCAATACCTCGAAATTAACAGAG 535
151 AspLeuValAsnThrSerValTyrAsnLLeuHisGlnGluAspArgLy 167
      |||
536 GACCTGGTTAAACCAAGTCTTTACAAATATCTTACATGAAGAAAGACAGAA 585
167 sAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrpT 184
      |||
586 GGATTTTCTTAAGAAATTTACAAATCTACAGTAAATGAGATTCTCTGGA 635
184 hTrAsnGluProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200
      |||
636 CAAATGAGACCCCAAGACAAAAAGCCATACATTTAATTTGCCGTATGTTG 685
201 MetLysThrProHisAspLLeLeuGlnAspLLeAsnAlaSerProGluMe 217
      |||
686 ATGAAGAAACACACATGATATTTCTGGAAGACATRAAACGCCAGCTCGAAT 735
217 tArgGlnArgTyrGluThrMetGlnCysPheAlaLeuSerGlnProArgA 234
      |||
736 GCGCCAGAGATATGAACAATCAGAGCTTCCCTGCTCAGCACAGAG 785
234 lAsMetMetGlnGlnGlnLysAspLeuGlnSerCysMetLLeCysValAla 250
      |||
786 CTATGATGGAGGAGGAGGAAATTTGCAATCTTGATGATCTGTGTGGCA 835
251 ArgArgLLeThrThrGlyGluArgThrPheProSerAsnProGluSerPh 267
      |||
836 CCGCCGATTTACTACAGAGCAAGAAACATTTCCATCAAAACCTCAGAGCTT 885
267 eLLeThrArgHisAspLeuSerGlyLysValValAsnLLeAspThrAsnS 284
      |||
886 TATTAACAGACATGATCTTTACGAAAGGTGTCAATATATAGTACAAAT 935
284 eTrLeuArgSerSerMetArgProGlyPheGlnAspLLeLLeArgArgCys 300
      |||
```

936 CACTGAGATTCCTCATGAGGCTGGCTTGCAAGATATATATCCGAGGCT 985
 301 ILeGlnArgPhePheSerLeuAsnAspGlyGlnSerTrpSerGlnLys 317
 986 ATTCAGAGATTTTATAGTAAATGATGGGCACTGATGTCACGAAACG 1035
 317 gHsTrpGlnGluAlaTrpLeuAsnGlnHsIslaGlnThrProValTrp 334
 1036 TCACATCAAGAAAGCTTATCTTATATGGCCATGCAGAAACCCAGTATATC 1085
 334 rPheSerLeuAlaAspGlyThrIleValThrAlaGlnThrLysSerLys 350
 1086 GATTCTCGTGGTGATGGAACATAAGTACTGACAGACGACAAAGCAA 1135
 351 LeuPheArgAsnProValThrAsnAspArgHsGlyPheValSerThrHs 367
 1136 CTCTTCGGAATCCTGTAAACAAATGATCCGACATGGCTTTGTCAACCCA 1185
 367 sPheLeuGlnArgGluGlnAsnGlyTrpArgProAsnProAsnProValG 384
 1186 CTTCCTTCAGAGAAACAGAAATGATATAGACCAAAACCCAAATCTGTTG 1235
 384 LysGlnGlyIleArgProProMetAlaGlyCysAsnSerSerValGlyGly 400
 1236 GACAAGGAGATTAGACCACTATGCTGATGCAACAGTTGCTGAGCGGC 1285
 401 MetSerMetSerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTrp 417
 1286 ATGAGTATGTCGCCAAACCAAGGCTTACAGATCCGACGACGAGGGCTTA 1335
 417 rGlyLeuAlaAspProSerThrThrGlyGlnMetSerGlyAlaArgTrpG 434
 1336 TGGCTTGCGCAGACCTAGCACACAGGCGCATGAGGAGCTAGGATG 1385
 434 LysLysSerSerAsnIleAlaSerLeuThrProGlyProGlyMetGlnSer 450
 1386 GGGGTTCCAGTAAACATAGCTTACCTTGACCCCTGGCGACGATGCAATCA 1435
 451 ProSerSerTrpGlnAsnAsnAsnTrpGlyLeuAsnMetSerSerProTrp 467
 1436 CCATCTTCCACAGAACCAACATAGGCTCAACATGATGATGACCCCCC 1485
 467 oHsGlySerProGlyLeuAlaProAsnGlnGlnAsnIleMetIleSerP 484
 1486 ACATGGGAGTCCGTGCTTCCGCCAAACGACAGAAATATCATGATTTCTC 1535
 484 rOArgAsnArgGlySerProLysIleAlaSerHsGlnPheSerProVal 500
 1536 CTCGTAATCGTGGAGTCCAAAGATAGCCCTCACATCAGTTTCTCCGT 1585
 501 AlaGlyValHsSerProMetAlaSerSerGlyAsnThrGlnLysHsSer 517
 1586 GCAGGtGTGCACTCCCATGSCATCTTGGCAATATCTGGAAACCCAG 1635
 517 rPheSerSerSerSerLeuSerAlaLeuGlnAlaIleSerGlyGlyValG 534
 1636 CTTTTCAGAGGCTCTCTCAGTGCCCGCAGACCATCAGTAAGAGTGTG 1685
 534 LysThrSerLeuLeuSerThrLeuSerSerProGlyProLysLeuAspAsn 550
 1686 GGACTTCCTCTTTATCTACTCTCATACACAGGCCCCCAATATGGATTAAC 1735
 551 SerProAsnMetAsnIleThrGlnProSerLysValSerAsnGlnAspSe 567
 1736 TCTCCCAATATGAAATATACCAACCAAGTAAGTAAGTAAGTAAGTTC 1785
 567 LysSerProLeuGlyPheTrpCysAspGlnAsnProValGlnSerSerM 584
 1786 CAAGAGTCCCTCTGGCTTTTATTGCGACCAAAATCCAGTGGAGATTCAA 1835
 584 eLysGlnSerAsnSerArgAspHsLeuSerAspLysGlnSerLysGln 600
 1836 TGTGTAGTCAAAATAGCAGATCACCTCAGTGCACAAAGAAATTAAGGAG 1885

601 SerSerValGlnGlyAlaGlnAsnGlnArgGlyProLeuGlnSerLysGln 617
 1886 AGCAGTGTAGGGGCGAGAAATCAAGGGGTCTTTGGAAAGCAAGG 1935
 617 yHsLysLysLeuLeuGlnLeuLeuThrCysSerSerAspArgGlyHs 634
 1936 TCATAAAAAATTACTGCAGTACTTACTCTCTCTGATGAGACGGGGTCT 1985
 634 LysSerSerLeuThrAsnSerProLeuAspSerSerCysLysGlnSerSer 650
 1986 ATTCCTCTTGACCAACTCCCGCTAGATCAAGTTGTAAAGATCTTCT 2035
 651 ValSerValThrSerProSerGlyValSerSerSerThrSerLysGlyVal 667
 2036 GTTAGTGTACACAGCCCTCTGGAAGTCTCTCTCATCATCTGGAGAGT 2085
 667 LysSerThrSerAsnMetHsGlySerLeuLeuGlnGlnLysHsArgI 684
 2086 ATCTCTACATCCAAATATGATGATGGTCTGTACAAAGAAACACCGCA 2135
 684 LeuLeuHsLysLeuLeuGlnAsnGlyAsnSerProAlaGlnValAlaLys 700
 2136 TTTTGCACAAGTTGCTGCAGAAATGGGAATTCACACGCTGAGTACGCAAG 2185
 701 IleThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysGln 717
 2186 ATTAAGTGCACAAAGCACTGGGAAAGACACGACAGATTAATTAATCTTGTGG 2235
 717 YAspGlyAsnValValLysGlnGlnLeuLeuSerProLysLysGlnVal 734
 2236 GAGCGAAATGTTGTCAAGCAGAGAGCGATAGTCTTAATAAAGAGAGA 2285
 734 sAsnAlaLeuLeuAlaArgTrpLeuLeuAspArgAspProSerAspAla 750
 2286 ATATAGCACTTCTTACATCTCTGTCAGACGAGATGATCTCTAGTATGCA 2335
 751 LeuSerLysGlnLeuGlnProGlnValGlnGlyValAspAsnLysMetSe 767
 2336 CTCTCTAAAGAACTACAGCCCAAGTGGAAGAGTGAGCAATTAATATGAG 2385
 767 rGlnCysThrSerSerThrIleProSerSerSerGlnGlnLysAspProL 784
 2386 TCAGTGCACAGCTCCACCATCTCTACTACTCAAGTAAGAAAGACCCCTA 2435
 784 YsIleLysThrGlnThrSerGlnGlnGlySerGlyAspLeuAspAsnLeu 800
 2436 AAATTAAGACAGAGACAAAGTAAGAGGATCTGAGACTTGGAATATCTTA 2485
 801 AspAlaIleLeuGlnLysLeuThrSerSerAspPheTrpAsnAsnSerI 817
 2486 GATGCTATTTCTTGATGTGATCTGACTAGTCTGACTTTTACATATATTCAT 2535
 817 eSerSerAsnGlySerHsLysLeuGlyThrLysGlnGlnValPheGlnGly 834
 2536 ATCTCTAAATGATGATCATCTGGGGAATAAGCAAGGTGTTTCAAGGA 2585
 834 hAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleArgPro 850
 2586 CTAAATCTCTGGGTTTGAAAAGTTTCAAGTCTGACGTATATTCGCCCT 2635
 851 ProThrAsnArgAlaValSerLeuAspSerProValSerValGlySerSe 867
 2636 CCATATTAACCGAGCGAGTCTCTGATAGCCCTGTTTCTTGGCTCAAG 2685
 867 rProProValLysAsnIleSerAlaPheProMetLeuProLysGlnProm 884
 2686 TCTCTCCAGTAAAAAATATTCAGTGTCTTCCCATGTTACCAAGAACCA 2735
 884 eLeuGlnGlyAsnProArgMetMetAspSerGlnGlnLysTrpGlySer 900
 2736 TGTGGGTGGGAATCCAAAGAAATGATGATGTCAGGAAATTAATGAGCTCA 2785

```

901 SermetGlyGlyProAsnArgAsnValThrValThrGlnThrProSerSe 917
|||||
2786 AGTATGGGTGGGCCCAACCGAATGTGACTGTGACTGACACTCTCTCCTC 2835
917 rGlyAspTrpGlyLeuProAsnSerLysAlaGlyArgMetGluProMetA 934
|||||
2836 AGGAGACTGGGGCTTACCAACTCAAGACCGGCGAGAAATGGAACCTTANGA 2885
934 snSerAsnSerMetGlyArgProGlyLysAspTyrAsnThrSerLeuPro 950
|||||
2886 ATTCAAACTCCATGGGAGACCGAGAGAGATTATTAATCTCTTACC 2935
951 ArgProAlaLeuGlyGlySerIleProThrLeuProLeuArgSerAsnSe 967
|||||
2936 AGACCTGCACGGGTGGCTTATTCACATTCGCTCTTGGTCTAAATAG 2985
967 rIleProGlyAlaArgProValLeuGlnGlnGlnGlnMetLeuGlnM 984
|||||
2986 CATACCAGGTGGAGACAGTATTCACACAGCAGCAGCAGATGCTTCAA 3035
984 eArgProGlyGlyIleProMetGlyMetGlyAlaAsnProTyrGlyGln 1000
|||||
3036 TGAGGCTGTGTAATCCCATGGGAAATGGGGCTAATCCCTATGGCCAA 3085
1001 AlaAlaAlaSerAsnGlnLeuGlySerTrpProAspGlyMetLeuSerMe 1017
|||||
3086 GCAGCAGCATCTAACCAACTGGGTTCTGGCCCGATGGCATGTGTCCAT 3135
1017 tGlnGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerL 1034
|||||
3136 GGAAACAATTTCTCATGGCAGCTCAAAATAGGCCCTTCTTAAAGAAATTC 3185
1034 euAspAspLeuValGlyProProSerAsnLeuGlyGlnSerAspGlu 1050
|||||
3186 TGATGATCTTGTGGGCCACTTCCAACTGGAAAGCCAGAGTACGAA 3235
1051 ArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaTh 1067
|||||
3236 AGAGCATTTATGGACACAGCTGCACACTCTTCTCACCAACACAGATGGCG 3285
1067 rGlyLeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnG 1084
|||||
3286 AGGCTGGAGAAATTTGACAGAGCTTTGGCACTTCCTGAACTTGTCAATC 3335
1084 InGlyGlnAlaLeuGluProLysGlnAspAlaPheGlnGlyGlnGluA 1100
|||||
3336 AGGACAGGCAATTAGACCCAAACAGATGCTTCCAAAGGCCAAGAAGA 3385
1101 AlaValMetMetAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAl 1117
|||||
3386 GCAGTAATGATGGATCAGAAAGCAGGATTTATTTGACAGACATACCAGC 3435
1117 agInGlyProProMetGlnGlyGlyPheHisLeuGlnGlnGlnSerProS 1134
|||||
3436 ACAGGGGCTCCAAATGCAAGAGGCTTTCAITTTCAAGGACATACACCAT 3485
1134 erPheAsnSerMetMetLeuAsnGlnMetLeuGlnGlnGlyAsnPheProLeu 1150
|||||
3486 CTTTAACTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 3535
1151 GlnGlyMetHisProArgAlaAsnIleMetArgProArgTyrAsnThrPr 1167
|||||
3536 CAAGGAATGACCCACGAGCCAACTCATGAGAACCCGACAAACACCC 3585
1167 oLysGlnLeuArgMetGlnLeuGlnGlnArgLeuGlnGlnGlnGlnPheL 1184
|||||
3586 CAAGCAACTTAAATGACACTTACAGCAGAGGCTGCAAGGGCCAGCATTTT 3635
1184 euAsnGlnSerArgGlnAlaLeuGlnLeuLysMetGlnLysAsnProThrAla 1200
|||||
3636 TGAATCAGAGCGGACAGGCACTTGAATTTGAAATGAAATGAAATGAAATG 3685
1201 GlyGlyAlaAlaValMetArgProMetMetGlnProGlnGlnGlyPheLe 1217

```

```

|||||
3686 GGTGTGCTGCGGTATGATGAGGCTTATGATGACAGCCAGAGGGTTTCT 3735
1217 uAsnAlaGlnMetValAlaGlnArgSerArgGluLeuLeuSerHisLSP 1234
|||||
3736 TAATGCTCAAAATGGTGGCCCAACGACAGAGAGAGTGTGAATGTCATCCT 3785
1234 hearGlnGlnIleArgValAlaMetMetMetGlnGlnGlnGlnGlnGln 1250
|||||
3786 TCCGACAAACAGAGGGTGGCTATGATGATGATGATGATGATGATGATGATG 3835
1251 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1267
|||||
3836 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3885
1267 nThrGlnAlaPheSerProProProAsnValThrLalaSerProSerMetA 1284
|||||
3886 AACCCAGGCTTTCAGCCCACTCTTAATGTGACTGCTTCCCGCAGCATGG 3935
1284 spGlyLeuLeuAlaGlyProThrMetProGlnAlaProProGlnGlnPhe 1300
|||||
3936 ATGGGCTTTTGGCAGACCCCAATGCCACAAAGCTCTCCGCAACAGTTT 3985
1301 ProTyrGlnProAsnTyrGlyMetGlyGlnGlnProAspProAlaPheG 1317
|||||
3986 CCATATCAACCAATTTATGAAATGGAGACAAACACAGATCCAGCCTTGG 4035
1317 yArgValSerSerProProAsnAlaMetMetSerSerArgMetGlyProS 1334
|||||
4036 TCGAGTGTCTAGTCTCCCAATGCAATGATGTGTGCAAGAAATGGGTCTCT 4085
1334 erGlnAsnProMetMetGlnHisProGlnAlaAlaSerLerTyrGlnSer 1350
|||||
4086 CCCAGAAATCCCATGATGCAACAGCCGAGGCTGCATTCATCATCATGATCC 4135
1351 SerGluMetLysGlyTyrProSerGlyAsnLeuAlaArgAsnSerSerPh 1367
|||||
4136 TCAGAAATGAAAGGCTGGCCATCAGAGAAATTTGGCCAGAACAGCTCTT 4185
1367 eSerGlnGlnGlnPheAlaHisGlnGlyAsnProAlaValTyrSerMetV 1384
|||||
4186 TTCCACAGCAGCATTTGCCACACAGGCAATCTCAGATGATATATG 4235
1384 aHisMetAsnGlySerSerGlyHisMetGlyGlnMetAsnMetAsnPro 1400
|||||
4236 TCCACATGAAATGACAGCATGGTGTACATGGGACAGATGAATGAACCC 4285
1401 MetProMetSerGlyMetProMetGlyProAspGlnLysTyrCys 1415
|||||
4286 ATGCCATGTCTGGCATGCTTAATGGTCTCTGATCAGAAATTAATCTGC 4330
seq_name: /cgn2_6/ptodata/1/pna/us091_COMB.seq.us-09-125-635-1
seq_documentation_block:
: Sequence 1, Application US/09125635
: GENERAL INFORMATION:
: APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE
: TITLE OF INVENTION: AIB1, A novel steroid receptor co-activator
: FILE REFERENCE: 49944
: CURRENT APPLICATION NUMBER: US/09/125,635
: PRIOR FILING DATE: 1998-08-21
: PRIOR APPLICATION NUMBER: 60/049,728
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 6835
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (201)..(4463)
US-09-125-635-1

```


alignment_scores:

Quality: 7352.50 Length: 1420
 Ratio: 5.200 Gaps: 1
 Percent Similarity: 99.577 Percent Identity: 99.366

alignment_block:

US-09-041-994-2 x US-09-125-635-1 ..

Align seg 1/1 to: US-09-125-635-1 from: 1 to: 6835

```

1 MetSerGlyLeuGlyGluAsnLeuAspProLeuAlaSerAspArgly 17
  |||||
201 ATGAGTGGATTAGAGAAACTTGATGATCCACTGGCCAGTATTCACGAAA 250
  |||||
17 sArgLeuLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGly 34
  |||||
251 ACGCAATTCGATGATGATCCAGGACAAAGCTTAACTGACAGTGTG 300
  |||||
34 IuysArgArgArgGluGlnGluSerIysTyrIleGluGluLeuAlaGlu 50
  |||||
301 AAAAAGCGAGCGGAGACGAGAAAGTAAATATATGAAAGATGGCTGAG 350
  |||||
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValIysProAs 67
  |||||
351 CTGATATCTGCCATCTTAGTATGATGACAAATTCATGTCACCAACGAG 400
  |||||
67 pIysCysAlaIleLeuIysGluThrValArgGlnIleArgGlnIleIysG 84
  |||||
401 TAAATGTCGATTTTAAAGAAACAGTAAGACAGATACGTCAATTAAGA 450
  |||||
84 IuGlnGlyIysThrIleSerAsnAspAspValGlnIysAlaAspVal 100
  |||||
451 AGCAGAGAAATCTATTTCCATATGATGATGATGATGATGATGATGATG 500
  |||||
101 SerSerThrGlyGlnGlyValIleAspIysAspSerLeuGlyProLeuLe 117
  |||||
501 TCTTCTACAGGCGAGGAGTATGATTAAGACTCTTAGACCGCTTTT 550
  |||||
117 uLeuGlnAlaLeuAspGlyPheLeuPheValAlaAsnArgGlnAlaAsnI 134
  |||||
551 ACTTCAGGCAATGGATGATGATGATGATGATGATGATGATGATGATG 600
  |||||
134 IeValPheValSerGluAsnValIleThrGlnIleGlnIleGlnGlnIle 150
  |||||
601 TTGATTTGATTCAGAAATGCTCACAAATCTGCAATATATACAGAGAG 650
  |||||
151 AspleuValAsnThrSerValTyrAsnIleuHsGluGluAspArgly 167
  |||||
651 GACCTGGTTACACAGAGTATTCATATCTTACATGAGAGAGACAGAAA 700
  |||||
167 sAspPheLeuIysAsnLeuProIysSerThrValAsnGlyValSerTyrP 184
  |||||
701 GGATTTCTTACAGAAATTCACAAATCTACAGTATATGAGATTCCTGGA 750
  |||||
184 hAsnGluProGlnArgGlnIysSerHisThrPheAsnCysArgMetLeu 200
  |||||
751 CAAATGAGACCCAAAGACAAATAATCATATTAATTCCTGATGTTG 800
  |||||
201 MetIysThrProHisAspIleLeuGluAspIleAsnAlaSerProGluMe 217
  |||||
801 ATGAAACACCCATGATATTCGAGACATTAACGCACTCCGAAAT 850
  |||||
217 IArgGlnArgTyrGluThrMetGlnCysPheAlaLeuSerGlnProArgA 234
  |||||
851 GCGCAGAGATATGAACAATGAGTGTGCTGCTGCTGCTGCTGCTGCTG 900
  |||||
234 IamMetGluGluGlyIuAspLeuGlnSerCysMetIleCysValAla 250
  |||||
901 CTATGATGAGAGAGGAGGAGATTTGCAATCTTATGATGATCTGTGTGCA 950
  |||||
251 ArgArgIleThrThrGlnGluArgThrPheProSerAsnProGluSerPh 267
  |||||

```

```

951 CGCCGATTACTACAGAGAAACATTTCCATCAACCCCTGAGACTT 1000
  |||||
267 eIleThrArgHisAspLeuSerGlyIysValValAlaAsnIleAspThrAsn 284
  |||||
1001 TATTACAGACATGATCTTTCAGGAAAGGTTGTCMAATATGATCAAAAT 1050
  |||||
284 eIleuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300
  |||||
1051 CACTGAGATCTCCATGAGGCTGCTTTCAGATATATATCCGAGGTGT 1100
  |||||
301 IlegIlnArgPhePheSerLeuAsnAspGlyIleInSerTyrPserGlnIysAr 317
  |||||
1101 ATTCAGAGATTTTATGATTAATGATGGAGAGCATGATGCTCCAGAAACG 1150
  |||||
317 gHisTyrGlnGluAlaTyrIleuAsnGlyHisAlaGluThrProValIysA 334
  |||||
1151 TCACATATCAGAAAGCTTATCTTAATGGCCATGAGAAACCCAGATATATC 1200
  |||||
334 rPheSerLeuAlaAspGlyThrIleValThrAlaGlnThrIysSerIys 350
  |||||
1201 GATTCGTGTTGCTGATGAGACTATGATGACTGCACAGACAAAGACAA 1250
  |||||
351 LeuPheArgAsnProValThrAsnAspArgHisGlyPheValSerThrHI 367
  |||||
1251 CTCTTCGAAATCTGTAACAAATGATGACATGATGCTTGTCTCAACCA 1300
  |||||
367 sPheLeuGlnArgGluGlnAsnGlyTyrArgProAsnProAsnProValG 384
  |||||
1301 CTCTCTTCAGAGAACAGATGATATGACCAACCCCAATCTGTTG 1350
  |||||
384 IyGlnGlyIleArgProPheMetAlaGlyCysAsnSerSerValGlyGly 400
  |||||
1351 GACAAGGATTAAGACCACTATGCTGATGATCAACAGTCTGGTAGGGG 1400
  |||||
401 MetSerMetSerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTyr 417
  |||||
1401 ATGAGTATGTCGCCAACCAAGGCTTACAGATGCGCAGACAGACAGGCTA 1450
  |||||
417 rGlyLeuAlaAspProSerThrThrGlyGlnMetSerGlyValAlaArgTyrG 434
  |||||
1451 TGCGTTGGCAGACCCCTAGCACACAGGCGAGATGAGTGGAGCTAGTATG 1500
  |||||
434 IyGlySerSerAsnIleAlaSerLeuThrProGlyProGlyMetGlnSer 450
  |||||
1501 GGGGTTCCAGTAACTAGCTTCAATGACCCCTGGGCCAGGATGCAATCA 1550
  |||||
451 ProSerSerTyrGlnAsnAsnAsnTyrGlyLeuAsnMetSerSerProPr 467
  |||||
1551 CCATCTTCTTACCAAGAACACACATATGAGGCTCAACATGATGATACCCC 1600
  |||||
467 oHisGlySerProGlyLeuAlaProAsnGlnGlnAsnIleMetIleSerP 484
  |||||
1601 ACATGGGAGGTCCTGCTGCTGCCCCAACACAGACAGATATGATTTCTC 1650
  |||||
484 rArgAsnArgGlySerProIysIleAlaSerHisGlnPheSerProVal 500
  |||||
1651 CTCGTAATCGGGAGTCCAAAGATAGCTCACAATGATTTTCTCCGTT 1700
  |||||
501 AlaGlyValHisSerProMetAlaSerSerGlyAsnThrGlyAsnHisSe 517
  |||||
1701 GCAGGTGTCATCTCCCATGCACTTTCTGCAATATGAGGAACACAGAG 1750
  |||||
517 rPheSerSerSerSerLeuSerAlaLeuGlnAlaIleSerGluGlyValG 534
  |||||
1751 CTTTTCAGAGCTCTCTCACTGAGCCCTGCAAGCCATGAGTAAGGTGAG 1800
  |||||
534 IyThrSerLeuLeuSerThrLeuSerSerProGlyProIysLeuAspAsn 550
  |||||
1801 GGACTTCCCTTTTATCTACTCTGTCATACACAGGCCCCAAATTTGGAATAC 1850
  |||||
551 SerProAsnMetAsnIleThrGlnProSerIysValSerAsnGlnAspSe 567
  |||||

```

1851 TCCTCCATATATGATATTTACCAACCAAGTAAAGTAAGCAATACAGANTTC 1900
567 rLysSerProLeuGlyPheTyrcysAspGlnAsnProValGluSerSerM 584
1901 CAAGAGTCCCTGTGGGCTTTATTGGACCAAAATCCAGTGGAGAGTTCAA 1950
584 eLcysGlnSerAsnSerArgAspHisLeuSerAspLysGluSerLysGlu 600
1951 TGTGTGAGTCAATATGACAGAGATCACCTCCAGTACAGAAAGAAAGTAAAGG 2000
601 SerSerValGluGlyAlaGluAsnGlnArgGlyProLeuGluSerLysGlu 617
2001 AGCAGTGTGGAGGGGCGACAGAAATCAAAAGGCTCTTTGGAAAGCAAGG 2050
617 yHisLysLysLeuLeuGlnLeuLeuThrcysSerSerAspAspArgGlyH 634
2051 TCATATAAAATTAAGTACGAGTACTTACCTGTTCTTGATGACCGGGGTC 2100
634 lAsSerSerLeuThrAsnSerProLeuAspSerSerCysLysGluSerSer 650
2101 ATTCCCTCTTGACCAACTCCCTCCAGATTCAAAGTTGTAAGAAATCTTCT 2150
651 ValSerValThrSerProSerGlyValSerSerSerThrSerGlyGlyVal 667
2151 GTTAACTGTACCAAGCCCTCTGAGTCTCTCTTACATCTGGAGAGGT 2200
667 lSerSerThrSerAsnMetHisGlySerLeuLeuGlnGlyLysHisArgI 684
2201 ATCCCTACATCCAAATATGCATGGGTCACTGTACMAAGAAAGCAACCGGA 2250
684 lLeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAlaLys 700
2251 TTTTGCACAAAGTTGCTGCAGAAATGGAAATTCACACAGCTGAGTAGCCAA 2300
701 lIleThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysG 717
2301 ATTACTGCAGAAAGCCACTGGAAAGACACACAGATTAATCTTCTTGTCG 2350
717 yAspGlyAsnValValLysGlnGluGlnLeuSerProLysLysGlyAla 734
2351 GGACGGAATTTGTTCMAAGCAGACAGCTTAAGTCTTAAGAAAGAGGAGA 2400
734 snAsnAlaLeuLeuArgTyrlLeuLeuAspArgAspAspProSerAspAla 750
2401 ATAAATGCACTTCTTAGATACCTGCTGGACAGGAGATCTTAGTGAATGCA 2450
751 LeuSerLysGluLeuGlnProGlnValGluGlyValAspAsnLysMetSe 767
2451 CTCTCTAAAGAACTACAGCCCAAGTGGAAAGAGTGAATTAATGAATGAG 2500
767 rGlnCysThrSerSerThrIleProSerSerSerGlnGlyLysAspProL 784
2501 TCAGTGCACACCTCCACCATCTCTAGCTCAAGTCAAGCAAGAAAGCCCTTA 2550
784 ySileLysThrGluThrSerGluGlyLysSerGlyAspLeuAspAsnLeu 800
2551 AAATTAAGCAGAGACAGTGAAGAGGATCTGGAGACTTGGATATATCTTA 2600
801 AspAlaIleLeuGlyAspLeuThrSerSerAspPheTyrcysAsnSerIle 817
2601 GATGCTATTCTGTGGATCTGACTAGTTCTGACTTTTCAATTAATTCAT 2650
817 eSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValPheGlnGlyT 834
2651 ATCTTCAATGCTAGTCACTCTGGGACTTAAGCAACAGGTGTTCAAGGAA 2700
834 hAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleArgPro 850
2701 CTAAATCTCTGTGGTTGAAGATTCACAGTCTGTCAAGTCTATTCGTCT 2750
851 ProTyrcysArgAlaValSerLeuAspSerProValSerValGlySerSe 867
2751 CCATATAACCGAGAGTCTCTGATAGCCCTGTTTCTGTGGCTCAAG 2800

867 rProProValLysAsnIleSerAlaPheProMetLeuProLysGlnProM 884
2801 TCCTCCAGTAAATAATATCAGTGTCTTCCCATGTGTACCAAGCAACCA 2850
884 eLLeuGlyLysProAsnMetAspSerGlnGluAsnTyrlGlySer 900
2851 TGTGTGGGTGGAAATCCAGAAATGATGATAGTACAGAAATTAATGCTCA 2900
901 SerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProSerSe 917
2901 AGTATGGGTGGGCCCAACCGAAATGTGACTGTGACTCAGACTCTTCTC 2950
917 rGlyAspTrpGlyLeuProAsnSerLysAlaGlyArgMetGluProMetA 934
2951 AGGAGACTGGGCTTACCAAACTCAAAAGCCGCGAAGATGAACTTAGA 3000
934 snSerAsnSerMetGlyArgProGlyGlyAspTyrcysThrSerLeuPro 950
3001 ATTCAAACTCCATGGGAAGACAGAGAGATTAATTAATCTTCTTACC 3050
951 ArgProAlaLeuGlyGlySerIleProThrLeuProLeuArgSerAsnSe 967
3051 AGACCTGCACCTGGGTGGCTCTATTCCACATTCCTCTTGGTCTAATAG 3100
967 rIleProGlyAlaArgProValLeuGlnGlnGlnGlnMetLeuGlnM 984
3101 CATACAGAGTGCAGACCCAGTATTCACACAGCAGCAGCAGATGCTTCAA 3150
984 eArgProGlyLysLysProMetGlyMetGlyAlaAsnProTyrlGlyGln 1000
3151 TGGGCTGTGTGAATCCCATGGGAATGGGGCTTAATCCCTATGGCCAA 3200
1001 AlaAlaAlaSerAsnGlnLeuGlySerTrpProAspGlyMetLeuSerMe 1017
3201 GCAGCAGCATTCACCAACTGGGTTCCTGGCCGATGGAGATGTTCAT 3250
1017 rGluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerL 1034
3251 GGAAACAAAGTTTCTCATGGCACATCAAAATAGGCCCTTCTTACGAATTC 3300
1034 euAspAspLeuValGlyProProSerAsnLeuGluGlyGlnSerAspGlu 1050
3301 TGGATGATTTGTGGGCCACCTTCCAACTCGAAAGCCAGAGTACGAA 3350
1051 ArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaTh 1067
3351 AGAGCATTTATGGACACAGCTGCACACTTCTCAGCAACACAGATGCCAC 3400
1067 rGlyLeuGlnGluIleAspArgAlaLeuGlyLysIleProGluLeuValAsnG 1084
3401 AGCCCTGGAAATTTGACAGAGCTTTGGGCATTCCTGAACTTGTCAATC 3450
1084 lGlnGlyAlaLeuGluProLysGlnAspAlaPheGlnGlyGlnGluAla 1100
3451 AGGGAAGCATTTAGAGCCCAACAGAGATCTTCCAAAGGCCAAAGAAAGA 3500
1101 AlaValMetMetAspGlnLysAlaGlyLeuTyrlGlyGlnThrTyProAl 1117
3501 GCAGTAATATGATGATCAGAGCAGAGATTAATGAGACAGACTTACCAGC 3550
1117 aGlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlnGlnSerProS 1134
3551 ACAGGGGCTCCCAATGCAAGAGGCTTTCATCTTCAGGAGCAATACACAT 3600
1134 ePheAsnSerMetMetAsnGlnMetAsnGlnGlnGlyAsnPheProLeu 1150
3601 CTTTTTAATCTATGATGAATTCAGATGAACCAAGCAAGCAATTTCTCTC 3650
1151 GlnGlyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPr 1167
3651 CAAGGAATGCAACCCAGAGCCCAACATCATGAGACCCCGGACAAACACCC 3700

```

1167 oLysGlnLeuArgMetGlnLeuGlnArgLeuGlnGlnGlnPheL 1184
|||||
3701 CAGGAACTTAGATGAGCTTCAGCAGAGGCTGCAGGGCAGCAGTTT 3750
|||||
1184 euASnGlnSerArgGlnAlaLeuGlnLeuLysMetGluAsnProThrAla 1200
|||||
3751 TGAATCAGAGCCGACAGGCACTGTAATGAAATGAAACCTACTGCT 3800
|||||
1201 GlyGlyAlaAlaValMetArgProMetMetGlnProGlnGlnPheLe 1217
|||||
3801 GGtGGtGGtGGtGGtGGtGGtGGtGGtGGtGGtGGtGGtGGtGGtGGt 3850
|||||
1217 uASnAlaGlnMetValAlaGlnArgSerArgGlnLeuLeuSerHisHisP 1234
|||||
3851 TAATGCTCAATGTGCTCCCAAGCAGCAGAGAGCTCTAATGATCATCT 3900
|||||
1234 hearGlnGlnArgValAlaMetMetMet.....GlnGln 1245
|||||
3901 TCCGACACAGAGGtGGtGGtGGtGGtGGtGGtGGtGGtGGtGGtGGtGGt 3950
|||||
1246 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1262
|||||
3951 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 4000
|||||
1262 ngGlnGlnGlnGlnGlnGlnGlnAlaPheSerProProAsnValThrA 1279
|||||
4001 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4050
|||||
1279 laSerProSerMetAspGlyLeuLeuAlaGlyProThrMetProGlnAla 1295
|||||
4051 CTTCCCCCAGCATGATGATGGCTTTGGCAGAGCCACCAATGCCAAGCT 4100
|||||
1296 ProProGlnGlnPheProTyrglnProAsnTyrglyMetGlyGlnInP 1312
|||||
4101 CCCTCCCAACAGTTTCCATATCAACCAAAATTGATGGAGACACCAACC 4150
|||||
1312 oaSPProAlaPheGlyArgValSerSerProProAsnAlaMetMetSers 1329
|||||
4151 AGATCCAGCCTTTGGTGGAGTGTCTAGTCTCTCCCAATGCATGATGTGCT 4200
|||||
1329 eaArgMetGlyProSerGlnAsnProMetMetGlnHisProGlnAlaAla 1345
|||||
4201 CAAGATGGGTCCTCCCAAGATCCCATGATGCACACCGCAGGCTGCA 4250
|||||
1346 SerTleTyrglnSerSerGlnMetLysGlyTrpProSerGlyLysLeuAl 1362
|||||
4251 TCCATCTATCTAGTCTCAGAAATGAAGGCTGCCATCAGGAAATTTTGGC 4300
|||||
1362 aArgAsnSerSerPheSerGlnGlnInPheAlaHisGlnGlyAsnProA 1379
|||||
4301 CAGGACAGCTCCTTTTCCAGCAGCAGAGTTGCCACACAGGGGAAATCTG 4350
|||||
1379 laValTySerMetValHisMetAsnGlySerSerGlyHisMetGlyGln 1395
|||||
4351 CAGTGTATATATAGTGTGACATGATGAGCAGAGTGTGATCATATGGGAG 4400
|||||
1396 MeASnMetAsnProMetProMetSerGlyMetProMetGlyProAspG 1412
|||||
4401 ATAAACATGAACCCCAAGCCCATGTCTGGCATGCTATGGGCTCTGATCA 4450
|||||
1412 nLysTyTyCys 1415
|||||
4451 GAATACTGC 4460

```

seq.name: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:US-60-068-511-1

seq.documentation_block:

; Sequence 1, Application US/60068511

; GENERAL INFORMATION:

; APPLICANT: Suen, Chen-Shian

; APPLICANT: Fraitl, Donald E.

; APPLICANT: Lytle, Richard C.

; TITLE OF INVENTION: Cloning and Expression of a Nuclear

```

; TITLE OF INVENTION: Receptor Coactivator Proteins and Uses Thereof
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Home Products Corporation
; STREET: One Campus Drive
; CITY: Parsippany
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA: Patent Release #1.0, Version #1.30
; APPLICATION NUMBER: US/60/068,511
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Walsh, Andrea C.
; REGISTRATION NUMBER: 34,988
; REFERENCE/DOCKET NUMBER: 97243-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 973-683-4117
; TELEFAX: 973-683-4117
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4789 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 185..4750
; US-60-068-511-1

```

alignment_scores:

Quality	Ratio	Length
7341.00	5.192	1423
Percent Similarity: 99.368	Percent Identity: 99.157	

alignment_block:

US-09-041-994-2 x US-60-068-511-1 ..

Align seg 1/1 to: US-60-068-511-1 from: 1 to: 4789

```

1 MetSerGlyLeuGlnGlnGlnLeuAspProLeuAlaSerAspSerArgly 17
|||||
185 ATGAGTGGATAGGAGAAACCTTGATCCAGTGGCCAGTGATTCACAGAAA 234
|||||
17 sarGlyLysLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGly 34
|||||
235 ACCGCAATTCACATGATCTCCAGACAAAGTCTTACCTGACGAGTGCTG 284
|||||
34 LuLysArgArgArgGlnGlnGlnSerLysTyrlleGlnGlnAlaGln 50
|||||
285 AAAAAGCAGACGAGCAGCAGGAAAGTAAATATATGAGAGATTGGCTGAG 334
|||||
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67
|||||
335 CTGATATCTGCCAATCTTACTGATATGACAAATTTCAATGTCAAACACAG 384
|||||
67 pLysCysAlaIleLeuLysGlnThrValAlaGlnIleArgGlnIleLysG 84
|||||
385 TAATGTGCGATTTTAAAGGAAACAGTAAGACAGATACGTCGAATTAAG 434
|||||
84 LuGlnGlyLysThrIleSerAsnAspAspValGlnLysAlaAspVal 100

```

435 AGCAAGGAAAACTATTTCATGATGATGATGTTCAAAAAGCCGATGTA 484
 101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeuLe 117
 485 TCTTCTACAGGGCAGGAGGATTTATGATAAGACTCTTAGAGCCGCTTTT 534
 117 uLeuGlnAlaLeuAspGlyPheLeuPheValValAsnArgGluAlaAsnI 134
 535 ACTTCAGGCAATTGGATGGTTTCCATTGTGTGTGTAATCGAGACGGAAAA 584
 134 lValAlaPheValSerGluAsnValThrGlnItyrLeuGlnItyrLysGln 150
 585 TTGTATTTTGTATCAGAAAATGTCACACAAATACCTGCAAATATAGCAAG 634
 151 AsPheValAsnThrSerValItyrAsnIleLeuHisGluGluAspArgLy 167
 635 GACCTGGTTAACACAAAGTTTACAAATATCTTACATGAAGAAGACAGAAA 684
 167 sAsPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTPT 184
 685 GGATTTTCTTAAGAATTACCAAAATCTACAGTTAATGAGATTCTCGGA 734
 184 hrAsnGluProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200
 735 CAAATGACACCCAAAGACAAAAAGCCATACCTTAATTCGCCGTATGTGG 784
 201 MetLysThrProHisAspIleLeuGluAspIleAsnAlaSerProGluMe 217
 785 ATGAATAACACACATGATATTCTGGAAAGACATTAACGCCAGTCTGAAT 834
 217 LArgGlnArgTyrGluThrMetGlnCysPheAlaLeuSerGlnProAla 234
 835 GCGCCAGAGATGTAACAATGCAAGTCTTGGCCCTGTCTCAGCCACAG 884
 234 lAmeMetGluGluGlyLysAspleuGlnSerCysMetIleCysValAla 250
 885 CTATGATGAGGAAGGGAAGATTGGCAATCTGTATGATCTGTGTGCA 934
 251 ArgArgIleThrThrGlyGluArgThrPheProSerAsnProGluSerPh 267
 935 CGCGCATTTACTACAGAGAAAAGAACATTTCCATCAAAACCTGAGAGCTT 984
 267 eLleThrArgHisAspleuSerGlyLysValValAsnIleAspThrAsn 284
 985 TATTACCGACATGATCTTTCAGAAAAGTTGTCAATATAGTACAAATT 1034
 284 eLleuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300
 1035 CACTGAGATCTCTCCATGAGGCTGGCTTTGAAGATATTAATCCGAAGGTGT 1084
 301 lIleGlnArgPhePheSerLeuAsnAspGlyLysItyrPheSerGlnLysAr 317
 1085 ATTCAAGATTTTATTAGTCTAAATGATGGGAGTCATGTCGCCAAGAACG 1134
 317 gHisTyrGlnGluAlaTyrLeuAsnGlyHisAlaGluThrProValTyrA 334
 1135 TCACATATCAAAAGCTATCTTAATAGGCCATGCAAAAACCCACAGATATTC 1184
 334 rPheSerLeuAlaAspGlyThrIleValIThrAlaGlnThrLysSerLys 350
 1185 GATTCCTGCTTGGCTGATGAACTATAGTACTGCCACAGACAAAAGCAAA 1234
 1235 CTCTTCGAAATCTGTAAACAATGATGACATGGCTTTGTCTCAACCCA 1284
 351 LeuPheArgAsnProValIThrAsnAspArgHisGlyPheValSerThrHI 367
 367 sPheLeuGlnArgGluGlnAsnGlyTyrArgProAsnProAsnProValG 384
 1285 CTTCCTTCAGAGAAAGAAATGATATAGCAAAACCAATCTGTGTG 1334
 384 LysGlnGlyIleArgProProMetAlaGlyCysAsnSerSerValGlyGly 400

1335 GACAAAGGATTAGACCACTATGCTGGATGCCAACAGTTCGGTAGCGGC 1384
 401 MetSerMetSerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTy 417
 1385 ATGAGTATGTGGCCAAACCAAGGCTTACAGATGCCGACAGAGAGGCTTA 1434
 417 rGlyLeuAlaAspProSerThrThrGlyGlnMetSerGlyAlaArgTyrG 434
 1435 TGGCTTGGCAGACCCCTTACACACAGGAGCAGATGATGGAGCTAGATGTG 1484
 434 LysGlySerSerAsnIleAlaSerLeuThrProGlyLysMetGlnSer 450
 1485 GGGGTTCCAGTACATAGCTTATGATGACCCCTGGGCGCAGGCTGCAATCA 1534
 451 ProSerSerTyrGlnAsnAsnAsnTyrGlyLeuAsnMetSerSerProP 467
 1535 CCATCTTCTCTACACAGACAAACAATATGGGTCACAAATGATAGGCCCCC 1584
 467 OHsGlySerProGlyLeuAlaProAsnGlnGlnAsnIleMetIleSerP 484
 1585 ACATGGAGTCTCTGGTCTTGGCCCCAAACAGCAGAAATATCATGATTTCTC 1634
 484 rOArgAsnArgGlySerProLysIleAlaSerHisGlnPheSerProVal 500
 1635 CTGTAATCGTGGAGTCCAAAGATAGCCTCACAATCACTTTTCTCTGTT 1684
 501 AlaGlyValHisSerProMetAlaSerSerGlyAsnThrGlyAsnHis 517
 1685 GCAGGTGTGCACTCTCCATGGCATCTTCTGGCAATACTGGGAACACAG 1734
 517 rPheSerSerSerSerLeuSerAlaLeuGlnAlaIleSerGlyGlyValG 534
 1735 CTTTTCCAGACGCTCTCTCAGTGGCCCTGCAAGCCATGAGTAAGTGTGG 1784
 534 LysThrSerLeuLeuSerThrLeuSerSerProGlyProLysLeuAspAsn 550
 1785 GGACTTCCCTTTTATCTACTCTGTCTATCACCAGGCCCCCAATTTGATTAAC 1834
 551 SerProAsnMetAsnIleThrGlnProSerLysValSerAsnGlnAspSe 567
 1835 TCTCCCAATATGAATATTAACCAACCAAGTAAGTAAGTAAGTACAGATTC 1884
 567 rLysSerProLeuGlyPheTyrCysAspGlnAsnProValGluSerSerM 584
 1885 CAAGAGTCTCTGGGCTTTTATGCGACCAAAATCCAACTGAGAGATTCAA 1934
 584 eCysGlnSerAsnSerArgAspHisLeuSerAspLysGluSerLysGlu 600
 1935 TGTGTCACTCAAAATGCGAGATTCACCTCACTGACACAAAGAAATAGGAG 1984
 601 SerSerValGluGlyAlaGluAsnGlnArgGlyProLeuGluSerLysGI 617
 1985 AGCACTGTTGAGGGGCGAGAGATCAAAAGGGTCTTTGGAAGCAAAAGG 2034
 617 YHisLysLysLeuLeuGlnLeuLeuThrCysSerSerAspAspArgGlyH 634
 2035 TCATTAATAAATTACTGCACTTACTTACCTGTCTTCATATACACGGGGTTC 2084
 634 lSerSerLeuThrAsnSerProLeuAspSerSerCysLysGluSerSer 650
 2085 ATTCTCTCTTACCAACTCCGCCCTAGATTCAAGTTGTAAAGAAATCTTCT 2134
 651 ValSerValIThrSerProSerGlyValSerSerSerThrSerGlyGlyVal 667
 2135 GTTAGTGTACACAGCCCTCTGAGAGTCCCTCTCTATATCTGGAGAGT 2184
 667 lSerSerThrSerAsnMetHisGlySerLeuLeuGlnGluLysHisArgT 684
 2185 ATCCCTATCATCCAAATATGCAATGGTCACTGTACAAAGAGAACCCGGA 2234
 684 lLeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAlaLys 700
 2235 TTTTGCACAAAGTTGCTGCGAATGGGAATTCACCAAGCTGAGGTAGCCAAAG 2284

701 IleThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysG1 717
 2285 ATTTACTGAGAGACCTGGGAAAGACACCCAGCTATTAACCTTCTGGTG 2334
 717 yAspGlyAsnValValLysGlnGlnLeuSerProLysLysGlnA 734
 2335 GGACGGAAATGTTGTCAMGAGACGACCTAAGTCTTAAGAAAGAGGA 2384
 734 snAsnAlaLeuLeuArgTyrLeuLeuAspArgAspProSerAspAla 750
 2385 ATAATGCACTTCTTGAATACCTGTCGACAGGATGATCCTAGTATCA 2434
 751 LeuSerLysGlnLeuGlnProGlnValGlnGlyAlaAspAsnLysMetSe 767
 2435 CTCTCTAAGAGACTACAGCCCAAGTGGAGAGAGGATTAATAAATGAG 2484
 767 rGlnCysThrSerSerThrIleProSerSerSerGlnGlnLysAspProL 784
 2485 TCACTGGACCAAGCTCCACCATTTCTAGCTCAAGTCAAGAGAAAGCCCTA 2534
 784 ySrlLeuLysThrGlnThrSerGlnGlySerGlyAspLeuAspAsnLeu 800
 2535 AAATTAAGACAGACACAAGTGAAGAGGATCTGGAGACTTGATTAATCTA 2584
 801 AspAlaIleLeuGlyAspLeuThrSerSerAspPheTyrAsnAsnSerIle 817
 2585 GATGCTATTTCTGGTGTGATCTGACTAGTCTGACTTTTACAAATAATCAT 2634
 817 eSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValPheGlnGlyT 834
 2635 ATCTCAATAGTAGTATCTGTGGGACTAACAACAGGTGTTCAAGGAA 2684
 834 hAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleArgPro 850
 2685 CTAATTTCTCTGGTGTGAAAGTTACACAGTCTGTCACACTTATTCGTCT 2734
 851 ProTyrAsnArgAlaValSerLeuAspSerProValSerValGlySerSe 867
 2735 CCAATTAACCCAGACAGTCTCTGATAGCCCTGTTTCTGTGGCTCAG 2784
 867 rProProValLysAsnIleSerAlaPheProMetLeuProLysGlnProM 884
 2785 TCTCCAGTAAATAATATACAGTGTTCCTCCATGTTACCAAGACCA 2834
 884 eLeuGlyGlyAsnProArgMetMetAspSerGlnGlnAsnTyrGlySer 900
 2835 TGTGGGTGGGAATCCAAAGATGATGATGATGAGGAAATTAATGCTCA 2884
 901 SerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProSerSe 917
 2885 AGTATGGGTGGGCCAAMCCGAATGTGACTGTGACTCAGACTCCTTCTCTC 2934
 917 rGlyAspTyrGlyLeuProAsnSerLysAlaGlyArgMetGlnProMetA 934
 2935 AGGAGACTGGGGCTTACCAACTCAAGCGCGCAAGATGGAACCTATGA 2984
 934 snSerAsnSerMetGlyArgProGlyGlyAspTyrAsnThrSerLeuPro 950
 2985 ATTCAAACTCCATGGGAGAGACAGAGAGATTAATAATCTTCTTACCC 3034
 951 ArgProAlaLeuGlyGlySerIleProThrLeuProLeuArgSerAsnSe 967
 3035 AGACTGTGACTGGGTGCTCTATTCACAGCTTGGCTCTTGGCTTAATAG 3084
 967 rIleProGlyAlaArgProValLeuGlnGlnGlnGlnMetLeuGlnM 984
 3085 CATACAGAGTGGAGACAGATTAATGCAACAGACAGACAGATGCTTCAAA 3134
 984 eTArgProGlyGlnIleProMetGlyMetGlyAlaAsnProTyrGlyGln 1000
 3135 TGAAGCCCTGGTGAATCCCAATGGGAATGGGGCTTAATCCTATGGCCAA 3184

1001 AlaAlaIleSerAsnGlnLeuGlySerThrProAspGlyMetLeuSerMe 1017
 3185 GCAGCAGCATTAACCAACTGGGTCTGGGCCCGAGTGGCATGTGTCCAT 3234
 1017 tGlnGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerI 1034
 3235 GGAAACAAGTTCTCTACAGCTCAAAATAGGCTCTTCTTGAAGATGCC 3284
 1034 euAspAspLeuValGlyProProSerAsnLeuGlyGlnGlnSerAspGlu 1050
 3285 TGGATGATCTTGTGGGCCACCTTCCAACTGGAAGGCCAAGAGACGAA 3334
 1051 ArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaTh 1067
 3335 AGAGCATTAATGAGACCAAGCTGCACACTTCTCAAGCAACACAGATGCCAC 3384
 1067 rGlyLeuGlnGlnIleAspArgAlaLeuGlyTleProGlnLeuValAsnG 1084
 3385 AGGCTGGAAACAATGTGACAGAGCTTGGCATTCCTCAACTTGTGCATC 3434
 1084 lngGlyAlaIleLeuGlnProLysGlnAspAlaPheGlnGlyGlnGlnA 1100
 3435 AGGACAGGCAATTAGAGCCCAACAGAGATGCTTCCAAAGCCCAAGAGCA 3484
 1101 AlaValMetMetAspGlnLysAlaGlyLeuTyrGlnThrTyrProAl 1117
 3485 GCAGTAATGATGATCGAAGCCAGGATTAATGAGACAGACTGCCACGC 3534
 1117 aglngIlyProProMetGlnGlyGlyPheHisLeuGlnGlnGlnSerProS 1134
 3535 ACAGGGGCTTCATTCGACAGAGGCTTTCATCTTCAGGAGCAATCACCAT 3584
 1134 erPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlnGlnGlnGlnGln 1150
 3585 CTTTAACTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 3634
 1151 GlnGlyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPr 1167
 3635 CAAGGAATGCCACCCAGACAGCCACATCATGAGACCCCGGAGAAACACCC 3684
 1167 OlyGlnLeuLeuArgMetGlnLeuGlnGlnArgLeuGlnGlnGlnPheL 1184
 3685 CAAGCAACTTGAATGACAGCTTTCAGACAGGCTCAGAGGCCACAGAGTTT 3734
 1184 euAsnGlnSerArgGlnAlaLeuGlnLeuLysMetGlnAsnProThrAla 1200
 3735 TGAATCAGAGCCGACAGGCACTGATTAATGAATAAGAAACCTACTGCT 3784
 1201 GlyGlyAlaAlaValMetArgProMetMetGlnProGln.....G1 1214
 3785 GGTGGTGTGGGTGATGAGGCTATGATGATGATGATGATGATGATGATGAT 3834
 1214 nGlyPheLeuAsnAlaGlnMetValAlaGlnArgSerArgGlnLeuLeuS 1231
 3835 GGGTTTCTTAATGCTCAAAATGTCGCCCAAGCAGCAGAGAGCTGCTAA 3884
 1231 erHisHisPheArgGlnGlnArgValAlaMetMetMet.....1243
 3885 GTCAATCACTTCCGACCAACAGAGGGGTGCTATGATGATGATGATGATGAT 3934
 1244 ...GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1259
 3935 CAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 3984
 1259 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1276
 3985 ACAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCTCTA 4034
 1276 snValThrAlaSerProSerMetAspGlyLeuLeuAlaGlyProThrMet 1292
 4035 ATGTGACTGCTTCCCAAGATGATGATGATGATGATGATGATGATGATGAT 4084
 1293 ProGlnAlaProProGlnGlnPheProTyrGlnProAsnTyrGlyMetG1 1309

```
|||||
4085 CCACAGCTCTCCGACACAGTTTCCATATCAACCAATTTATGGAATGGG 4134
1309 yGlnGlnProAspProAlaPheGlyValSerProProAlaAla 1326
4135 ACAACAACGAGATCCAGCCTTTGGTGCAGTGTCTAGTCCCTCCCAATGCAA 4184
1326 eNMeSerSerArgMetGlyProSerGlnAsnProMetMetGlnHisPro 1342
4185 TGATGTCGTCAGAAATGGGTCCCTCCAGAAATCCCATGATGCAACCCCG 4234
1343 GlnAlaAlaSerIleTyrGlnSerSerGlnMetGlyTyrProSerGln 1359
4235 CAGGCTGCATCCATCTATCAAGTCTCCAGAAATGAAGGCGTGGCCATCAGG 4284
1359 yAsnLeuAlaArgAsnSerSerPheSerGlnGlnPheAlaHisGln 1376
4285 AATTTGGCCAGGACAGCCTTTCCAGCAGCAGATTGGCCACACAGG 4334
1376 yAsnProAlaValTyrSerMetValHisMetAsnGlySerSerGlnHis 1392
4335 GGAATCCTCAGTGTATGATGTGTGCACATGATGACAGCAGTGGTGCAC 4384
1393 MetGlyGlnMetAsnMetAsnProMetProMetSerGlyMetProMetGln 1409
4385 ATGGGACATGACATGAACCCCATGCCCATGTGTGGCATGCTATGGG 4434
1409 yProAspGlnLysTyrCys 1415
4435 TCCTGATCAGAAATACTGC 4453

seq_name: /cgn2_6/ptodata/1/pna/US093_COMB.seq:us-09-359-922-3962

seq_documentation_block:
: Sequence 3962, Application US/09359922
: GENERAL INFORMATION:
: APPLICANT: Hyseq, Inc.
: TITLE OF INVENTION: NOVEL CONTFICS OBTAINED FROM VARIOUS CDNA
: TITLE OF INVENTION: LIBRARIES
: FILE REFERENCE: 20411-752CON1
: CURRENT APPLICATION NUMBER: US/09/359,922
: CURRENT FILING DATE: 1999-07-22
: EARLIER APPLICATION NUMBER: US 09/205,155
: EARLIER FILING DATE: 1998-12-03
: NUMBER OF SEQ ID NOS: 13203
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 3962
: LENGTH: 6855
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(6855)
: OTHER INFORMATION: n = A,T,C or G
US-09-359-922-3962

alignment_scores:
    Quality: 7340.50      Length: 1424
    Ratio: 5.191          Gaps: 2
    Percent Similarity: 99.298    Percent Identity: 99.087

alignment_block:
US-09-041-994-2 x US-09-359-922-3962 ..
Align seg 1/1 to: US-09-359-922-3962 from: 1 to: 6855

1 MetSerGlyLeuGlyGlnAsnLeuAspProLeuAlaSerAspSerArgly 17
|||||
201 ATGAGTGGATTGAGAGAAAATTGATTCACATGGCCAGTGAATTCCAGAAA 250
17 sArgLysLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGlyG 34
|||||
```

```
251 AGCAAAATGGCCATGTGATATCTCCAGACAAAGTCTTACCTGCAAGTGTG 300
34 LuLysArgArgArgGlnGlnGlnSerLysTyrIleGlnGlnLeuAlaGln 50
301 AAAAAGCGAGACGGAGAGAAAATTAATATATATGAGAAATGGCTGAG 350
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPhaAsnValLysProAs 67
351 CTGATATCTGCCCAATCTTATGTATATGACAAATTTCAATGTCAAACCAA 400
67 pLysCysAlaIleLeuLysGlnThrValArgGlnIleArgGlnIleLysG 84
401 TAAATGTGGCATTTTAAAGAAACAGTAAGACAGATACGTCAATATAAG 450
84 LuGlnGlyLysThrIleSerAsnAspAspAspValGlnLysAlaAspVal 100
451 AGCAAGAAAAACTATTTCCAAATGATGATGATGTCCAAAAGCCGATGTA 500
101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeuLe 117
501 TCTTCTACAGGCGCAGGAGTTATTTGATAAAGACTCCTTAGGACCCGCTT 550
117 uLeuGlnAlaLeuAspGlyPheLeuPheValValAsnArgGlnAlaAsnI 134
551 ACTTCAGGCAATGGATGGTTCCTATTTGGTGAATGACAGACGAAACA 600
134 LeValPheValSerGlnAsnValThrGlnTyrLeuGlnTyrLysGlnGln 150
601 TTGTATTTGATCAGAAATGTCAACAAATACCTGCAATATATAAGCAAG 650
151 AspLeuValAsnThrSerValTyrAsnIleLeuHisGlnGlnAspArgly 167
651 GACCTGTTAACACAAAGTGTTCAAATATCTTACATGAAGAAGACAGAAA 700
167 sAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTyrP 184
701 GGATTTCTTTAAGAAATTTACAAATCTACAGTTAATGAGTTTCCGTGA 750
184 hAsnGlnProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200
751 CAATATGAGACCCCAAGACAAAAGCCATATCAATTAATGGCGGTATGGT 800
201 MetLysThrProHisAspIleLeuGlnAspIleAsnAlaSerProGluMe 217
801 ATGAAAAACCAACATGATATTCGGAAGACATTAACGCCAGCCTGAAAT 850
217 tArgGlnArgTyrGlnThrMetGlnCysPheAlaLeuSerGlnProArgA 234
851 GCCCCAGACATATGAAACAATGCAGTGTTCCTGCTCAGCCACGAG 900
234 lAmMetGlnGlnGlyGlnAspLeuGlnSerCysMetIleCysValAla 250
901 CTATGATGGAGGAAGGGAAGATTGCAATCTGTATGATCTGTGGCA 950
251 ArgArgIleThrThrGlyGlnArgThrPheProSerAsnProGluSerh 267
951 CGCCGCAATTACTACAGAGAAAAGAACATTTCCATCAACCCCTGAGAGCT 1000
267 eIleThrArgHisAspLeuSerGlyLysValValAsnIleAspThrAsn 284
1001 TATTTACCAGACATGATCTTTCAGAGAAAGTTGTCATATATAGATCAAA 1050
284 eLeuArgSerSerMetArgProGlyPheGlnAspIleIleArgArgCys 300
1051 CACTGAGATCTCTCCATGAGGCGCTTGAAGATATTAATCCGAAGGTGT 1100
301 lIleGlnArgPhePheSerLeuAsnAspGlyGlnSerThrSerGlnLysAr 317
1101 ATTCAAGATTTTATTAATGCTAAATGATGGCAGTCATGCTCCACAGAAAG 1150
317 gHisTyrGlnGlnAlaTyrLeuAsnGlnLysIleAlaGlnThrProValTyr 334
1151 TCACATATCAAGAAAGCTTATCTTATATGCGCATGACAGAAACCCAGTATAT 1200
```

334 rppheserleuAlaaspGlyThrIleValThrAlaGlnThrLysSerLys 350
 |||||
 1201 GATTCTGGTGGCTGATGGAACTATAGTACTGTACACAGACAAAAGGAAA 1250
 351 LeuPheArgAsnProValThrAsnAspArgHisGlyPheValSerThrHis 367
 |||||
 1251 CTCTCCGAATCTCTGAACAATGATCGACATGGCTTTGTCTCAACCCA 1300
 367 sPheLeuGlnArgGluGlnAsnGlyTyrArgProAsnProAsnProValG 384
 |||||
 1301 CTTCCTTCAGAGAGAAACAGATGATATAGACCAACCAATCCGTGTG 1350
 384 lvgIneGlyIleArgProPromeTAlaGlyCysAsnSerSerValGlyGly 400
 |||||
 1351 GACAAAGGATATGACCACTGATGGTGGATGACAAAGTTCTGGTAGCGCG 1400
 401 MetSerMetSerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTyr 417
 |||||
 1401 ATGATGATGTCGCCCAACCAAGGCTTACAGATGCCAGACGAGGCGCTTA 1450
 417 rGlyLeuAlaAspProSerThrThrGlyGlnMetSerGlyAlaArgTyrG 434
 |||||
 1451 TGGCTTGGCAGACCTAGCACACAGGCGAGATGAGTAGGAGTATG 1500
 434 lvgLysSerSerAsnIleAlaSerLeuThrProGlyProGlyMetGlnSer 450
 |||||
 1501 GGGGTTCCAGTAACATAGCTCTCATTAAGCCCTGGCCAGCATGCAATCA 1550
 451 ProSerSerTyrGlnAsnAsnAsnTyrGlyLeuAsnMetSerSerProPr 467
 |||||
 1551 CATCTTCTCTACAGAACCAACACTATGGGCTCAACATGATGATGCCCCC 1600
 467 ohISGlySerProGlyLeuAlaProAsnGlnAsnIleMetIleSerP 484
 |||||
 1601 ACATGGGAGTCTCGGTCTTGGCCCCAAACGACAGAAATATCATGATTTC 1650
 484 roArgAsnArgGlySerProLysIleAlaSerHisGlnPheSerProVal 500
 |||||
 1651 CTGCTATGTCGGAGTCCAAAGATAGCCCTCACATCAAGTTTCTCCGTG 1700
 501 AlaGlyValHisSerProMetAlaSerSerGlyAsnThrGlyAsnHisS 517
 |||||
 1701 GCAGGTGTGACATCTCCATGGCATCTTTCGCAATACTGGGAACCAAG 1750
 517 rPheSerSerSerSerLeuSerAlaLeuGlnAlaIleSerGlyGlyValG 534
 |||||
 1751 CTTTTCCAGAGCTCTCTCAGTGCCTCGCAAGCATCAGTGAAGGTGTGG 1800
 534 lYThrSerLeuSerThrLeuSerSerProGlyProLysLeuAspAsn 550
 |||||
 1801 GGCATTCCTCTTATCTACTCTCTCATCACAGGCCCAATGGATTAAC 1850
 551 SerProAsnMetAsnIleThrGlnProSerLysValSerAsnGlnAspS 567
 |||||
 1851 TCTCCCAATATGAAATTAATCCCAACCAAGTAAGTAAGCAATGAGATTTC 1900
 567 rLysSerProLeuGlyPheTyrCysAspGlnAsnProValGluSerSerM 584
 |||||
 1901 CAAGAGTCTCTGTGGCTTTATATGACACCAAAATCCAGTGAAGATGTCAA 1950
 584 efcYsGlnSerAsnSerArgAspHisLeuSerAspLysGluSerLysGlu 600
 |||||
 1951 TGTGTAGTCAAAATAGCAGAGATCACCTCAGTACAAAGAAATGAAGAG 2000
 601 SerSerValGluGlyValGlnAsnGlnArgGlyProLeuGlnSerLysG 617
 |||||
 2001 AGCAGGTGTGAGGGGCAAGAAATCAAGGGTCTTTGGAAAGCAAAGG 2050
 617 yHisLysLysLeuLeuGlnLeuLeuThrCysSerSerAspAspArgGlyH 634
 |||||
 2051 TCATAAAAAAATTACTGACATTACTACTCTGTCTGTGATGACCGGGGTC 2100

634 lSserSerLeuThrAsnSerProLeuAspSerSerCysLysGluSerSer 650
 |||||
 2101 ATTCTCTCTTGACCAACTCCCCCTAGATTCAAGTTGTAAGATCTTCT 2150
 651 ValSerValThrSerProSerGlyValSerSerSerThrSerGlyGlyVal 667
 |||||
 2151 GTTAGTGTACCAAGCCCTCTGTGAGTCTCTCTCTACATCTGTGAGAGAT 2200
 667 lSerSerThrSerAsnMetHisGlySerLeuLeuGlnLysHisArgI 684
 |||||
 2201 ATCCTCTACATCCAAATATGATGATGGTCTACTGTTCACAGAGAAACCCGGA 2250
 684 lLeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGlyValAlaLys 700
 |||||
 2251 TTTTGCAACAAGTTGTGCAAAATGGGAATTCACACACTGAGTAGCCAG 2300
 701 lIeThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysG 717
 |||||
 2301 ATTACTGCAGAAACCACTGGGAAAGAACACAGCAGTATACTTCTGTGG 2350
 717 yAspGlyAsnValValLysGlnGluGlnLeuSerProLysLysGluAla 734
 |||||
 2351 GGACGGAAATGTTGTCAAGCAGAGCAGCTAAAGTCTTAAGAAAGAGAGA 2400
 734 sPAsnAlaLeuLeuArgTyrLeuLeuAspArgAspProSerAspAla 750
 |||||
 2401 ATATGCACCTCTTAGATATACCTCTGTGACAGGGATATCTTATGTATGA 2450
 751 LeuSerLysGluLeuGlnProGlnValGlyValAlaAspAsnLysMetS 767
 |||||
 2451 CTCTCTTAAGAACTACACAGCCCAAGTGAAGAGTGGATATTAATTAAG 2500
 767 rGlnCysThrSerSerThrIleProSerSerSerGlnGlyLysAspProI 784
 |||||
 2501 TCAGTGCACAGCTCCACCACTTCCTAGCTCAAGTCAAGAGAAAGACCTTA 2550
 784 ySrlLeuThrGluThrSerSerGluGluGlySerGlyAspLeuAspAsnLeu 800
 |||||
 2551 AAATTTAGACAGACAGCAAGTGAAGAGGATCTGGAGACTTGGATTAATCTTA 2600
 801 AspAlaIleLeuGlyAspLeuThrSerSerAspPheTyrAsnAsnSerI 817
 |||||
 2601 GATGCTATTCTTGATCTGATCTGACTAGTTCGACTTTACAAATTAATTCAT 2650
 817 eSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValPheGlnGlyT 834
 |||||
 2651 ATCCTCAAAAGTATGATCTCTGGGACTAAGCAACAGGTGTTTCAAGGAA 2700
 834 hrasnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleArgPro 850
 |||||
 2701 CTAAATCTCTGGGTTGAAAAGTTCAAGTCTGTGACAGTCTATTCTCTCT 2750
 851 ProTyrAsnArgAlaValSerLeuAspSerProValSerValGlySerS 867
 |||||
 2751 CCATATTAACCGACAGACTGTCTCGAATAGCCCTGTTCTGTGGCTCAAG 2800
 867 rProProValLysAsnIleSerAlaPheProMetLeuProLysGlnProm 884
 |||||
 2801 TCTCTCAGTAAAAAATATCAGTCTTCCCATGTTACCAAAACCAACCA 2850
 884 eHeuGlnGlyAsnProArgMetMetAspSerGlnGluAsnTyrGlySer 900
 |||||
 2851 TGTGGGTGGGAATCCAAAGATATGATGATGTCAGAAAAATTAATTAAGCTCA 2900
 901 SerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProSerS 917
 |||||
 2901 AGTATGGGTGGCCCAACCAAAATGTGACTGTGACTCAGACACTCTCTCTC 2950
 917 rGlyAspTrpGlyLeuProAsnSerLysAlaGlyArgMetGluPrometa 934
 |||||
 2951 AGGAGACTGGGGCTTACCAAACTCAAAAGCCGCGCAAGTAAGAACCTATGA 3000
 934 snSerAsnSerMetGlyArgProGlyGlyAspTyrAsnThrSerLeuPro 950


```

|||||
3001 ATTGAAATCCATGGAGACGACGAGAGATTAATATCTTCTTACC 3050
951 ArpProalaleuGlySerIleProthrleuProleuArgSerAsnSe 967
|||||
3051 AGACCTGCACGTGGGTGCTTATTCACATGCTCTTCGCTTAATAG 3100
967 rIleProGlyAlaArgProValLeuGlnGlnGlnImeLeuGlnm 984
|||||
3101 CATACAGAGTCCGAGACCAAGTATTGCAACAGACAGACAGATGCTTCAA 3150
984 etaArgProGlyGluIleProMetGlyMetGlyAlaAsnProTyrGlyIn 1000
|||||
3151 TGAGGCTGTGTAATCCCATGGGAATGGGGCTTAATCCCTATGGCCAA 3200
1001 AlAlaAlaSerAsnGlnLeuGlySerTyrProAspGlyMetLeuSerMe 1017
|||||
3201 GCACGACGATTTAACCAACATGGGTTCTCGGCCGATGGCATGTTCAT 3250
1017 tGluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerL 1034
|||||
3251 GGACCAAGATTCTCATGGCCTCAAAATAGGCTCTCTTAGAATTCCTCC 3300
1034 euAspAspLeuValGlyProProSerAsnLeuGlnGlyGlnSerAspGlu 1050
|||||
3301 TGGATGATCTTGTGGGCCACCTTCCAACTGGAAAGCCAGAGTACGAA 3350
1051 ArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaTh 1067
|||||
3351 AGAGATTATTTGACACACAGCTGACACCTCTCTCAGCAACACAGATGCC 3400
1067 rGlyLeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnG 1084
|||||
3401 AGGCTGTGAAATAATTGACAGAGCTTTGGCATTCCTGAACTGTGCAATC 3450
1084 lngGlyAlaLeuGlnIleProLysGlnAspAlaPheGlnGlyGlnIle 1100
|||||
3451 AGGCAAGGCTTTAGAGCCCAACAGAGTCTTCCAGGCCAAGAACCA 3500
1101 AlAlaValMetLeuAspGlnIleAlaGlyLeuTyrGlyInThrTyrProAl 1117
|||||
3501 GCAGTAATGATGATCAGACAGAGGATTAATGACAGACATACCCAGC 3550
1117 agGlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlyGlnSerProS 1134
|||||
3551 ACAGGGGCTCCCAATGCAAGAGGCTTTCATCTTCAGGAAATCACCAT 3600
1134 erPheAsnSerMetLeuAsnGlnImeLeuGlnGlnGlyAsnPheProLeu 1150
|||||
3601 CTTTAACTCTATGATGAAATCAGATGAACCAAGCAAGCAATTTTCTCTC 3650
1151 GlnGlyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPr 1167
|||||
3651 CAAGGAATGACCCACAGCCACATCATGAGACCCCGACAAACACACCC 3700
1167 olyGlnLeuArgMetGlnLeuGlnGlnArgLeuGlnGlyGlnGlnPheL 1184
|||||
3701 CAAGCAACTTGAATGACGCTTCAGCAGAGGCTGACAGGGCCAGCGTTT 3750
1184 euAsnGlnSerArgGlnAlaLeuGlnLeuLysMetGluAsnProThrAla 1200
|||||
3751 TGAAATCAGAGCGACAGGCACTGTAATTGAAAATGAAAACCTTACTCT 3800
1201 GlyGlyAlaAlaValMetArgProMetMetGlnPro.....G1 1213
|||||
3801 GGTGGTCTCGGTGATGAGGCTTATGATGACCCCGAGTGAGCTCCCA 3850
1213 nGlnGlyPheLeuAsnAlaGlnMetValAlaGlnArgSerArgGluLeuL 1230
|||||
3851 GCAGGGTTTCTTAATGCTCAAAATGTGCCCAACGACGAGAGGCTGC 3900
1230 euSerHisHisPheArgGlnGlnArgValAlaMetMetMet..... 1243
|||||

```

```

3901 TAACTCATCTCTCCGACACAGAGGCTGCTATGATGATGCACAGCAG 3950
1244 .....GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1258
3951 CAGCAGACGACAAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 4000
1258 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1275
4001 GCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4050
1275 roAsnValThrAlaSerProSerMetAspGlyLeuLeuAlaGlyProThr 1291
4051 CTAATGTACTGCTTCCCGACAGATGATGGGCTTTTGGAGGAGCCACA 4100
1292 MetProGlnAlaProProGlnGlnGlnPheProTyrGlnProAsnTyrGlyMe 1308
4101 ATGCCACAGCTCTCCGCAACAGTTTCCATATCACCAAATTTATGGAT 4150
1308 tGlyGlnGlnProAspProAlaPheGlyArgValSerSerProAsnAla 1325
4151 GGGACCAACACAGATCCAGCCTTGGTGCAGTGTCTAGTCTTCCCAATG 4200
1325 lAmMetSerSerArgMetGlyProSerGlnAsnProMetMetGlnHis 1341
4201 CAATGATGCTGTCAGAAATGGGTCCCTCCAGAAATCCCATGATGCACAC 4250
1342 ProGlnAlaAlaSerIleTyrGlnSerSerGlyMetLysGlyTyrProSe 1358
4251 CCGCAGGCTGCATCCATCTATCAGTCTCAGAAATGAAAGGCTGGCCATC 4300
1358 rGlyAsnLeuAlaArgAsnSerPheSerGlnGlnGlnPheAlaHisG 1375
4301 AGGAATTTGGCCAGGAACAGCTCTTTCACACAGCAGATTTCCGCCAC 4350
1375 lngGlyAsnProAlaValTyrSerMetValHisMetAsnGlySerSerGly 1391
4351 AGGGAACTCTGCAGTGTATGATGTGTCACATGAATGCCAGCAGTGT 4400
1392 HisMetGlyGlnMetAsnMetAsnProMetProMetSerGlyMetProMe 1408
4401 CACATGGGACAGATGACATGAAACCCCATGTCATGTCTGCATGCTCAT 4450
1408 tGlyProAspGlnLysTyrCys 1415
4451 GGGTCTGTGATCAGAAATCTGTC 4472

seq_name: /cgn2_5/prodata/1/pna/US093_COMB.seq:us-09-359-922-3962
seq_documentation_block:
; Sequence 3962, Application US/09359922A
; GENERAL INFORMATION:
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Liu, Jin
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-752CON1
; CURRENT APPLICATION NUMBER: US/09/359,922A
; CURRENT FILING DATE: 1999-07-22
; EARLIER APPLICATION NUMBER: US 09/205,155
; EARLIER FILING DATE: 1998-12-03
; EARLIER APPLICATION NUMBER: US 09/034,341
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 13203
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3962
; LENGTH: 6855
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(6855)
; OTHER INFORMATION: n = A,T,C or G
US-09-359-922-3962

```


alignment_scores:

Quality: 7340.50 Length: 1424
 Ratio: 5.191 Gaps: 2
 Percent Similarity: 99.298 Percent Identity: 99.087

alignment_block:

US-09-041-994-2 x US-09-359-922-3962 ..

Align seg 1/1 to: US-09-359-922-3962 from: 1 to: 6855

1 MetSerGlyLeuGlyGluAsnLeuAspProLeuAlaSerAspSerArgly 17
 |||||
 201 ATGAGTGATGAGGAAACCTTGATCCACTGCCAGTGATTCACGAAA 250
 17 sAtgLyLeuProCysAspThrProGlyGlyLeuThrCysSerGly 34
 |||||
 251 ACGCAATTCGATGTACTCCAGGACAAGCTTACCTGACGTGGTG 300
 34 LuysArgArgArgGluGlnGlnSerLysTyrIleGluGluLeuAlaGlu 50
 |||||
 301 AAAACCGAGACGGAGCGAGAAAGTAATATATTAAGAAATGGCTGAG 350
 51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67
 |||||
 351 CTGATATCTGCACATCTTAGTATGATGACATTTCAATGTCACAAACGAG 400
 67 pLyScysAlaIleLeuLysGluThrValArgGlnIleArgGlnIleLysG 84
 |||||
 401 TAAATGTCGATTTTAAAGAAACAGTAACAGATACGTCGAATTAAG 450
 84 LuGlnGlyLysThrIleSerAsnAspAspValGlnLysAlaAspVal 100
 |||||
 451 AGCAGAGAAATCTATTCCATGATGATGATGATGATGATGATGATGATG 500
 101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeu 117
 |||||
 501 TCCTCTACAGGCGAGGAGTATGATTAAGACCTTAGACCGCTTTT 550
 117 uLeuGlnAlaLeuAspGlyPheLeuPheValValAsnArgGluAlaAsnI 134
 |||||
 551 ACTTCAGGCAATGGATGGTTCTTCTATTTGGTGAATCGAGACGGAACA 600
 134 lValAlaPheValSerGluAsnValThrGlnTyrLeuGlnTyrLysGlnGlu 150
 |||||
 601 TTGATTTGATACAGAAATGTCACACATACCTGCAATATTAACCAAGAG 650
 151 AspLeuValAsnThrSerValTyrAsnIleLeuHisGluGluAspArgly 167
 |||||
 651 GACCTGGTTAACACAGTGTTCACATATCTTACATACAGACAGACAGAAA 700
 167 sAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrpT 184
 |||||
 701 GGATTTTCTTAAGAAATTTACCAAAATCTACAGTAAATGGAGTTTCCGGA 750
 184 hTrasnGluProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200
 |||||
 751 CAAATGAGACCCAAAGCAAAAGCCATCATTTAATTGGCCGTATGTGG 800
 201 MetLysThrProHisAspIleLeuGluAspIleAsnAlaSerProGlu 217
 |||||
 801 ATGAAAAACACACATGATATTTCTGGAAGACATTAACCCGAGTCTGAAAT 850
 217 tArgGlnArgTyrGluThrMetGlnCysPheAlaLeuSerGlnProArg 234
 |||||
 851 GCGCCAGAGATATGAACAATGACAGTCTTGGCCCTGCTCAGCAGCAGAG 900
 234 lMetMetGluGluGlyGluAspLeuGlnSerCysMetIleCysValAla 250
 |||||
 901 CTATGATGAGGAGGAGGAGGATTTGCAATCTTGTATGATCTGTGTGCA 950
 251 ArgArgIleThrThrGlyGluArgThrPheProSerAsnProGlnSerPh 267

951 CGCCGATTAATACAGGAGAAACATTTCCATCAACCAACCTGAGACTT 1000
 267 eIleThrArgHisAspLeuSerGlyLysValValAsnIleAspThrAsn 284
 |||||
 1001 TATTACAGACATGATCTTTCAGGAAAGTTGTCAAATATATACAAATT 1050
 284 eIleuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300
 |||||
 1051 CACTGAGATCTCCATGAGGCTGGCTTTGAAGATATATATCCGAAGGT 1100
 301 lIleGlnArgPhePheSerLeuAsnAspGlyGlnSerTrpSerGlnLysAr 317
 |||||
 1101 ATTACAGATTTTATGCTTAATGATGAGGAGCATGCTGCCAGAAACG 1150
 317 gHisTyrGlnGluAlaTyrLeuAsnGlyHisAlaGluThrProValTyrA 334
 |||||
 1151 TCACATACAGAAAGCTTATCTTAATGGCCATGACAGAAACCCAGTATATC 1200
 334 rPheSerLeuAlaAspGlyThrIleValThrAlaGlnThrLysSerLys 350
 |||||
 1201 GATTCCTGTTGGCTGATGAGAACTATGACTGACACAGACAAAAAGCAA 1250
 351 LeuPheArgAsnProValThrAsnAspArgHisGlyPheValSerThrH 367
 |||||
 1251 CTCTCCGAAATCCTGTAAACAAATGATGACATGCTTTGCTCAACCCA 1300
 367 sPheLeuGlnArgGluGlnAsnGlyTyrArgProAsnProAsnProValG 384
 |||||
 1301 CTTCCTTCAGAGACAGAAATGATGATAGACCAACCAATCTCTGTG 1350
 384 LyGlnGlyIleArgProProMetAlaGlyCysAsnSerSerValGlyGly 400
 |||||
 1351 GACAGGATTAAGACCACTATGGCTGTGATGCAACACTGTGGTGGCGGC 1400
 401 MetSerMetSerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTy 417
 |||||
 1401 ATGAGTATGTCGCCAACAACCAAGCTTACAGATGCCGAGACAGAGGCTA 1450
 417 rGlyLeuAlaAspProSerThrThrGlyGlnMetSerGlyValaArgTyrG 434
 |||||
 1451 TGGCTTGGACAGACCTTAGACACACAGGAGATGATGAGACTGTGATG 1500
 434 LyGlySerSerAsnIleAlaSerLeuThrProGlyProGlyMetGlnSer 450
 |||||
 1501 GGGGTTCCAGTAACATGATGATGACCTTGGCCAGGAGCATCAATCA 1550
 451 ProSerSerTyrGlnAsnAsnAsnTyrGlyLeuAsnMetSerSerProP 467
 |||||
 1551 CCATCTTCCTACCAAGAACCAACATATGGGCTCAACATGATGAGTACCC 1600
 467 oHisGlySerProGlyLeuAlaProAsnGlnGlnAsnIleMetLieserP 484
 |||||
 1601 ACATGGGAGTCTGCTGCTGGCCCAACAGACAAATATCATGATTTCTC 1650
 484 rArgAsnArgGlySerProLysIleAlaSerHisGlnPheSerProVal 500
 |||||
 1651 CTCGTATCGGGAGTCCAAAGATAGCTTCACATCATGATTTTCTCCGTT 1700
 501 AlaGlyValHisSerProMetAlaSerSerGlyLysnThrGlyAsnHis 517
 |||||
 1701 GCAGGTGCACTCTCCATGGCATTTCTTGCAATATCTGGGAACCCACAG 1750
 517 rPheSerSerSerSerLeuSerAlaLeuGlnAlaIleSerGlnGlyValG 534
 |||||
 1751 CTTTTCAGAGAGCTCTCTCAGTGCCCTGCAAGCATCATGAGAAAGTGTGG 1800
 534 LyThrSerLeuLeuSerThrLeuSerSerProGlyProLysLeuAspAsn 550
 |||||
 1801 GGACTTCCCTTTATCTACTCTGTCATCACCAGGCCCAATGAGATTAAC 1850
 551 SerProAsnMetAsnIleThrGlnProSerLysValSerAsnGlnAspSe 567
 |||||

1851 TCTCCCATATATGATTTATCCCAACCAAGTAAAGTAAGCAATCAGATTC 1900
567 rLysSerProLeuGlyPheTyCysAspGlnAsnProValGluSerSer 584
|||||
1901 CAAGAGTCTCTGGGCTTTATTGGACCAAAATCCAGTGAGAGTTCA 1950
584 eCysGlnSerAsnSerArgAspHisLeuSerAspLysGluSerLysGlu 600
|||||
1951 TGTGTACAGCAAAATAGCAGAGATCACCTCAGTACAAAGAAAGTAAGAG 2000
601 SerSerValGluGlyAlaGluAsnGlnArgGlyProLeuGluSerLysGlu 617
|||||
2001 AGCAGTGTGGAGGGGACAGAGATCAAGGGCTCTTTGGAAAGCAAGAG 2050
617 yHisLysLysLeuLeuGlnLeuLeuThrCysSerSerAspAspArgGly 634
|||||
2051 TCATATAAAATTAAGTACAGTACTTACTCTCTCTGTGATGACCGGGGTC 2100
634 sAspSerLeuThrAsnSerProLeuAspSerSerCysLysGluSerSer 650
|||||
2101 ATTCTCTCTTGACCAACTCCCTCAGATTCAAGTTGTAAAGAACTTCT 2150
651 ValSerValThrSerProSerGlyValSerSerSerThrSerGlyGly 667
|||||
2151 GTTAGTGTACCAAGCCCTCTGAGTCTCTCTCTACATCTGGAGAGAT 2200
667 sSerSerThrSerAsnMetHisGlySerLeuLeuGlnGluLysHisArg 684
|||||
2201 ATCCCTACATCCAAATATGCATGAGGTCACTGTACAAAGAAAGACCGGA 2250
684 LeuLeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAlaLys 700
|||||
2251 TTTTGCACAGATTGCTGCAGAAATGGAAATTCACAGCTGAGGTACCAAG 2300
701 IleThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysGlu 717
|||||
2301 ATTACTGCAGAGCCACTGGGAAAGACACACAGATATACTTCTTGTCG 2350
717 yAspGlyAsnValValLysGlnGluGlnLeuSerProLysLysGlyAla 734
|||||
2351 GGACGGAATGTTGTCAAGCAGACAGCTTAAGTCCCTTAAGAAAGAGAGA 2400
734 snAsnAlaLeuLeuArgTyrlLeuLeuAspArgAspAspSerSerAla 750
|||||
2401 ATAAAGCACTTCTTAGATACCTGTGTGACAGGAGATCTCTAGTAGTGA 2450
751 LeuSerLysGluLeuGlnProGlnValGluGlyValAspAsnLysMetSer 767
|||||
2451 CTCTCTAAAGAACTACAGCCCAAGTGGAGAGAGTGAATTAATAATGAG 2500
767 rGlnCysThrSerSerThrIleProSerSerSerGlnGluLysAspProL 784
|||||
2501 TCAGTGCACACAGCTCCACCATCTCTAGCTCAAGTCAAGAGAAAGACCTTA 2550
784 yHisLeuThrGlnThrSerGlnGluLysSerGlyAspLeuAspAsnLeu 800
|||||
2551 AATATTAAGACAGACAGAGTGAAGAGGATCTGGAGACTTGGATTAATCTTA 2600
801 AspAlaIleLeuGlyAspLeuThrSerSerAspPheTyAsnAsnSerIle 817
|||||
2601 GATGCTATTCTTGATCTGATCTGACTTCTTACAAATAATTCCT 2650
817 eSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValPheGlnGlyT 834
|||||
2651 ATCCCTCAATGCTACTCATCTGGGACTTAAGCAACAGGTGTTCAAGGAA 2700
834 hrAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleArgPro 850
|||||
2701 CTAATTCTCTGGGTTGAAGATTACAGTCTGTGACGCTATTGCTCT 2750
851 ProTyAsnArgAlaValSerLeuAspSerProValSerValGlySerSer 867
|||||
2751 CCATTAACCGAGCAGTGTCTGTGATAGCCCTGTTTCTGTGGCTCAAG 2800

867 rProProValLysAsnIleSerAlaPheProMetLeuProLysGlnPro 884
|||||
2801 TCTCTCAGTAAAAAATATCAGTGTCTTCCCAATGTTACCAAGCAACCCA 2850
884 eLeuGlyLysAsnProArgMetAspSerGlnGluAsnTyrlGlySer 900
|||||
2851 TGTGGGTGGGAATCCAAAGATGATGATGATGATGATGATGATGATGATGAT 2900
901 SerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProSerSer 917
|||||
2901 AGTATGGGTGGGCCAAACCGAAATGTGACTGTGACTGACTGACTGCTTCTTC 2950
917 rGlyAspTrpGlyLeuProAsnSerLysAlaGlyArgMetGluProMet 934
|||||
2951 AGGAGACTGGGCTTACCAAACTCAAAAGCCGCGAGAAATGGAACCTATGA 3000
934 snSerAsnSerMetGlyArgProGlyGlyAspTyrlAsnThrSerLeuPro 950
|||||
3001 ATTCAAACTCCATGGGAAAGACAGAGAGAGATTAATATCTTCTTTACCC 3050
951 ArgProAlaLeuGlyGlySerIleProThrLeuProLeuArgSerAsnSer 967
|||||
3051 AGACCTGCCTGGTGCTCTTATTCACATTCCTCTTGGCTTAAATG 3100
967 rIleProGlyAlaArgProValLeuGlnGlnGlnGlnMetLeuGln 984
|||||
3101 CATACAGGTGGCAGACCAAGTATGCAACAGCAGCAGCAGATGCTTCAAA 3150
984 eArgProGlyLysIleProMetGlyMetGlyAlaAsnProTyrlGlyGln 1000
|||||
3151 TGAGGCTGTGTGAATCCCATGGGAATGGGGCTTAATCCCTATGGCCA 3200
1001 AlaAlaAlaSerAsnGlnLeuGlySerTrpProAspGlyMetLeuSerMet 1017
|||||
3201 GCAGACGATCTTAACCAACTGGGTTCTTGGCCGATGCGATTTGTCCAT 3250
1017 rGluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerL 1034
|||||
3251 GGAACAAGTCTTCATGCGACATCAAAATAGGCCCTCTTCAAGAAATTC 3300
1034 euAspAspLeuValGlyProProSerAsnLeuGlnGlyGlnSerAspGlu 1050
|||||
3301 TGGATGATCTTGTGGCCACCTTCCAACTTGAAGGCGAGTAGACGA 3350
1051 ArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 1067
|||||
3351 AGAGCATTTTGGACACAGCTGCACACTCTTCTCAGCAACACAGATGCCAC 3400
1067 rGlyLeuGlnGluIleAspArgAlaLeuGlyTleProGluLeuValAsnG 1084
|||||
3401 AGCCTGGAAAGAAATGACAGAGCTTGGGCAATTCCTCAACTTGTCAATC 3450
1084 rGlnGlnAlaLeuGlnProLysGlnAspAlaPheGlnGlyGlnGluAla 1100
|||||
3451 AGGAGACGCGATTAAGAGCCCAAGAGATGCTTCCAAAGGCGCAAGAGGA 3500
1101 AlaValMetMetAspGlnLysAlaGlyLeuTyrlGlyGlnThrTyrlProAl 1117
|||||
3501 GCAGTAATATGATGATCAGAGCAGAGATTAATGAGACAGACTTACCAGC 3550
1117 aGlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlnGlnSerProS 1134
|||||
3551 ACAGGGGCTCCAAATGCAAGGAGGCTTTCATCTTCAGGAGCAATCACAT 3600
1134 ePheAsnSerMetMetAsnGlnMetAsnGlnGlnGlyAsnPheProLeu 1150
|||||
3601 CTTTAACTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3650
1151 GlnGlyMetHisProArgAlaAsnIleMetArgProArgTyrlAsnThrPr 1167
|||||
3651 CAAGGAATGCAACCCAGAGCCCAACATCATGAGACCCCGGACAAACACCCC 3700

```

1167 ouysglnleuargmetglnleunglnlnargleunglnglnglnphel 1184
|||||
3701 CAAGCACTTAGATGCAGCTTCAGACAGAGGCGTGCAGGGCCACACTTTT 3750
|||||
1184 euasnglnserargglnalaleuqluleuylsmetcluasnprrhrala 1200
|||||
3751 TGAATCAGAGCCGACAGCAGCTTGAATGAAAATGAAAACCTACTGCT 3800
|||||
1201 glyglyalalavalmetargprometmetglnpro.....gl 1213
|||||
3801 GGGGTGGTGGGGTATGAGGCGCTATGATGCACGCCAGGTGAGCTCCA 3850
|||||
1213 nglnnglypheleuasnalaglmetvalalaglnarsserargglnleul 1230
|||||
3851 GCAGGGTTTTCTTAATGCTCAAAATGCTCCGCCACAGCAGAGAGCTGC 3900
|||||
1230 euserhishispheargnglnlnarqvalalametmetc..... 1243
|||||
3901 TAAATCATCATCTCCGACAACAGAGGGTGGCTATGATGCAGCAGCAG 3950
|||||
1244 .....glnnglnglnglnglnglnglnglnglnglnglnglngl 1258
|||||
3951 CAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 4000
|||||
1258 nglnnglnglnglnglnglnglnglnglnthrglnalapheserprop 1275
|||||
4001 GCACACGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4050
|||||
1275 roasnvalthralsarprosermetaspolyleuvalaglyprothr 1291
|||||
4051 CTAAATGTGACTGCTTCCCGACAGATGGATGGGCTTTTGGCAGAGCACA 4100
|||||
1292 metproglinalarproglnglnlnpheprotyrlnproasntlycylme 1308
|||||
4101 ATGCCCAAGCTCTCCGACACAGTTTCATATCAACCAAAATTATGAAAT 4150
|||||
1308 tclnglnglnproaspproalaphelgylarqvalaserpropasna 1325
|||||
4151 GGGACACACACAGATCCAGCTTTGGTGGAGTGTAGTCTCCCAATG 4200
|||||
1325 lametwetserserargmetgylproserglnasnprometmetglnhis 1341
|||||
4201 CAATGATGTGCTCAAGAAATGGGTCCTCCAGAAATCCCATGATGCAAC 4250
|||||
1342 proglinalalaserlletyrlnlnserserglnmetlycyltrprose 1358
|||||
4251 CCGCAGGCTGCATCCATATCATGCTCGAAATGAAGGCTGGCCATC 4300
|||||
1358 rglasnleualalargasnseraserpheserglnglnglnphelalhisg 1375
|||||
4301 AGGAAATTTGGCCAGACAGCTCTTTCCCGACAGAGTTGGCCACG 4350
|||||
1375 lnglyasnproalavaltyrsermetvalhismetasnnglysersergly 1391
|||||
4351 AGGGGAAATCTGCAGTGTATGATGTGCACATGAATGACAGCAGTGG 4400
|||||
1392 hismetgylglnmetasnmetasnprometprometserglnmetrome 1408
|||||
4401 CACATGGGACAGATGAACATGAACCCCATGCCATGTGGCATGGCTAT 4450
|||||
1408 tclproaspglnlytyrcys 1415
|||||
4451 GGGTCTCTGATCAGAAATACTGC 4472

```

seq_name: /cgn2_6/prodata/1/pna/US6017_COMB.seq:US-60-172-373-13551

```

; seq_documentation_block:
; Sequence 13551, Application us/60172373
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using

```

```

; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polym
; FILE REFERENCE: GX-0006 P
; CURRENT APPLICATION NUMBER: US/60/172,373
; CURRENT FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 25,772
; SOFTWARE: PERL Program
; SEQ ID NO 13551
; LENGTH: 7956
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 44586.3
US-60-172-373-13551

```

alignment_scores:

```

Quality: 7324.50      Length: 1424
Ratio: 5.184          Gaps: 3
Percent Similarity: 99.228      Percent Identity: 99.017

```

alignment_block:

US-09-041-994-2 x US-60-172-373-13551 ..

Align seg 1/1 to: US-60-172-373-13551 from: 1 to: 7956

```

1 metserglyleunglyluasnleuaspproleualaseraspserrgly 17
|||||
216 ATAGTGGATTTAGAGAAAACTTGATCCAGTGGCCAGTATTCACAGAAA 265
|||||
17 sarglylsleuprocyasprthproglnglnglyleuthrcyssergly 34
|||||
266 ACCGAAATGGCCATGTGATCTCAGGACAGAGGTCTTACCTGCAAGTGG 315
|||||
34 lufsarqarqarggluglngluserlystrlleglugluleualaglu 50
|||||
316 AAAAAGCGAGCGGAGCGAGGAAAGTAATATTTGAAGATTGGCTGAG 365
|||||
51 leuilserrlaasnleuserasprlleaspsnphasnvallysproas 67
|||||
366 CTGATATCTGCCCAATCTTAGTATTTGACAAATTCATCAATCCAAAC 415
|||||
67 plvscysalalleuylsgluthrvalarqglnllearglnlleargln 84
|||||
416 TAAATGGGATTTTAAAGAAACAGTAAGACAGATACGTAATATAAG 465
|||||
84 lueglnlystrhileaserasnspaspyalalnuysalaaaspyal 100
|||||
466 AGCAAGGAAAACTATTTCCAATGATGATGATGTTCAAAAGCCGATGTA 515
|||||
101 serserthrglynglnlyvalleaspylaspserserleuglyproleule 117
|||||
516 TCTTCTACAGGGCGAGGAGTATTGATTAAGACTCCTTAGGACCGCTTT 565
|||||
117 uleuglnalaleuaspglypheleuphevalalasnargglnalalasn 134
|||||
566 ACTTCAGGCAATGGATGGTTCCTATTTGGTGAATCGAGAGGAAACA 615
|||||
134 levalphevalsergluasvalthrglnlytleuglnlntlylsglnglu 150
|||||
616 TTGTATTTGTATCAGAAAATGTACACAAATATCTGCAATATAAGCAAG 665
|||||
151 aspleuvalasnthrservaltyrasnilleuunhsglucluaspargly 167
|||||
666 GACCTGTTAACACAAAGTGTTCATATCTTACATGAGAGAGACAGAAA 715
|||||
167 sasppheleuylsasleuuprolyserthryalaasnglyvalsertrpt 184
|||||
716 GGATTTTCTTAAGAAATTTACAAAATCTACAGTTAATGAGATTTCTGGA 765
|||||
184 hrasngluproglnarqglnlysserhsthrpheasnrcysargmetleu 200
|||||
766 CAATATAGACCCAAAGACAAAAAGCCATACATTTAATGGCGGTATGTTG 815

```

201 MetIysThrProHisAspIleLeuGluAspIleAsnAlaSerProGluMe 217
|||||
816 ATGAAACACACACATGATATCTCGAAGACATAAACGCCACTCCTGAAT 865
217 TargGlnArgTyrGluThrMetGlnCysPheAlaLeuSerGlnProArg 234
|||||
866 GCGCCACAGATATGAAACAAATGACGTCTTGCCCTGTCTCAGCCACGAG 915
234 IamEtmGluGluGlyGluAspLeuGlnSerCysMetIleCysValAla 250
|||||
916 CTATGATGGAGGAGGGAAGATTTCGATCTTGTATGATCTGTGTGCA 965
251 ArgArgIleThrThrGlyGluArgThrPheProSerAsnProGluSerP 267
966 CGCCGATTTACTACAGGAGAAACATTTCCATCAAAACCTGAGAGCTT 1015
267 eIleThrArgHisAspLeuSerGlyLysValValAsnIleAspThrAsn 284
|||||
1016 TATTACCGACATGATCTTTCAGGAAAGGTGTCATATAGATTCAAAATT 1065
284 erLeuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300
|||||
1066 CACGAGATCCTCCATGAGGCGCTTGGAATATATATCCGAAGGTGT 1115
301 IleGlnArgPhePheSerLeuAsnAspGlyGlnSerTrpSerGlnLysAr 317
|||||
1116 ATTCAGAGATTTTATGCTAAATGATGGGCGATGTGTCAGAAACG 1165
317 gHisTyrGlnGluAlaTyrLeuAsnGlyHisAlaGluThrProValTyr 334
|||||
1166 TCACCTATCAGAACCTTATCTTAATGGCCATGCAGAAACCCCATATATC 1215
334 rghSerLeuAlaAspGlyThrIleValThrAlaGlnThrLysSerLys 350
|||||
1216 GATTCGTGGTGGTGAAGACATAGTATGATGCACAGACAAAAAGCAA 1265
351 LeuPheArgAsnProValThrAsnAspArgHisGlyPheValSerThrH 367
|||||
1266 CTCTTCGAAATCCTGTAAACAAATGATCGACATGGCTTGTCTCAACCA 1315
367 sPheLeuGlnArgGluGlnAsnGlyTyrArgProAsnProAsnProValG 384
|||||
1316 CTTCCTTCAGAGACACAGATGATATAGACCAACCAATCCTGTGTG 1365
384 IyngIyGlyIleArgProProweAlaGlyCysAsnSerSerValGly 400
|||||
1366 GACAAGGAGATTAGACCACTATGGCTGATGCACAGTTGGTGGGCGC 1415
401 MetSerMetSerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTy 417
|||||
1416 ATGAGTATGTGCGCAACCAAGGCTTACAGATGCCGAGCAGGCGCTTA 1465
417 rGlyLeuAlaAspProSerThrThrGlyGlnMetSerGlyAlaArgTyr 434
|||||
1466 TGGCTTGGCAACCTTACACACACAGGCGACATAGTGGACTAGGTATG 1515
434 IyngIySerSerAsnIleAlaSerLeuThrProGlyProGlyMetGlnSer 450
|||||
1516 GGGGTTCAGTAACATAGCTTATGAAACCTGGGCGAGCATGCATCA 1565
451 ProSerSerTyrGlnAsnAsnAsnTyrGlyLeuAsnMetSerSerProP 467
|||||
1566 CCATCTTCCTACACAGAACCAACATATGGGCTCAACATGATAGCCCCC 1615
467 oHisGlySerProGlyLeuAlaProAsnGlnGlnAsnIleMetIleSerP 484
|||||
1616 ACATGGGAGTCTGTGCTTGGCCCAACACACAGAAATATCATGATTTCTC 1665
484 roArgAsnArgGlySerProLysIleAlaSerHisGlnPheSerProVal 500
|||||
1666 CTGTAATTCGTGGGAGTCCAAAGATAGCTCACATCAGTTTTCCTGCTT 1715

501 AlaGlyValHisSerProMetAlaSerSerGlyAsnThrGlyAsnHisSe 517
|||||
1716 GCAGGTGGCACTCTCCATGCGATCTTGTGGCAATCTGGGAACACAG 1765
517 rPheSerSerSerLeuSerAlaLeuGlnAlaIleSerGluGlyValG 534
|||||
1766 CTTTCCAGCAGCTCTCTCAGTGGCTGCAAGCAGCATGTAAGGTGG 1815
534 IyThrSerLeuLeuSerThrLeuSerSerProGlyProLysLeuAspAsn 550
|||||
1816 GGACTTCCTCTTATCTACCTGTGCATCAGCAGGCCCAAAATGTGATTAAC 1865
551 SerProAsnMetAsnIleThrGlnProSerLysValSerAsnGlnAspSe 567
1866 TCTCCATATATGAATATATCCCAACCAAGTAAAGTAAAGCAATTCAGATTC 1915
567 rLysSerProLeuGlyPheTyrCysAspGlnAsnProValGluSerSer 584
|||||
1916 CAAGAGTCTCTGGGCTTTTATTCGACCAAAATCCAGTGGAGAGTTCAA 1965
584 eCysGlnSerAsnSerArgAspHisLeuSerAspLysGluSerLysGlu 600
|||||
1966 TGTGTCACTCAATAGCAGAGATCACCTCAGTGAACAAAGAAAGTAAGAG 2015
601 SerSerValGluGlyAlaGluAsnGlnArgGlyProLeuGluSerLysG 617
|||||
2016 AGCAGGTGTGAGGGGCGAGAAATCAAAAGGGTCTTGTGAAACCAAAAG 2065
617 yHisLysLysLeuLeuGlnLeuLeuThrCysSerSerAspAspArgGlyH 634
|||||
2066 TCATRAAAATTAATGCGATTACTTACTCTTCTTCTGATATACCGGGGTC 2115
634 IIsSerSerLeuThrAsnSerProLeuAspSerSerCysLysGluSerSer 650
|||||
2116 ATTCCTCCTTGACCAACTCCCCCTAGATTCAGATGTGAAGAAATCTCT 2165
651 ValSerValThrSerProSerGlyValSerSerSerThrSerGlyGly 667
|||||
2166 GTTAGGTGCACACACCCCTCGAGGTCCCTCTCATATCTGAGGAGT 2215
667 IIsSerSerThrSerAsnMetHisGlySerLeuLeuGlnGluLysHisArg 684
|||||
2216 ATCTCTTACATCCAAATATGATGGGTCTACTTTCACAGACAAACGACCG 2265
684 IeLeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAlaLys 700
|||||
2266 TTTTGCAACAAGTGTGCGAGAATGGAAATTCACACAGTGAAGTGAAGCA 2315
701 IleThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysG 717
|||||
2316 ATTACTGCAGAAAGCCACTGGGAAAGACACCGACAGTATTAACCTTCTTGG 2365
717 yAspGlyAsnValValLysGlnGlnGlnLeuSerProLysLysLysGlu 734
|||||
2366 GGACGGAAATTTGTCAAGACGAGACAGACTTAAGTCTTAAGAAAGAGAGA 2415
734 snAsnAlaLeuLeuArgTyrIleLeuAspArgAspAspProSerAspAla 750
|||||
2416 ATTAATGCACCTCTTAGAATACCTGCTGGACAGGATGATCTTAAGATGCA 2465
751 LeuSerLysGluLeuGlnProGlnValGluGlyValAspAsnLysMetSe 767
|||||
2466 CTCTCTTAAGAACTTACAGCCCAAGTGGAAAGAGTGGCAATTAAGAGAG 2515
767 rGlnCysThrSerSerThrIleProSerSerSerGlnGluLysAspProL 784
|||||
2516 TCAGTGCACCGCTCCACCATTTCTAGCTCAAGTCAAGACAGAAACCCCTA 2565
784 ysiIeLysThrGluThrSerGluGlnGlySerGlyAspLeuAspAsnLeu 800
|||||
2566 AAATTAAGACAGACAAAGTGAAGGGGATCTGGAAGACTTGATTAATCTA 2615
801 AspAlaIleLeuGlyAspLeuThrSerSerAspPheTyrAsnAsnSerII 817

2616 GATGCAATTCCTTGATCTGACTAGTCTGACTTTTACAAATATTCAT 2665
 817 eSerSerasnGlySerHisLeuGlyThrIysGlnGlnValPheGlnGlyT 834
 2666 ATCCCTCAAAATGAGTATCATCTGGGGAGCTAAGCAACAGGTGTTCAAGAA 2715
 834 hrAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleaaGPro 830
 2716 CTAATTCCTCGGGTTTGAAAAAGTTTCACAGCTGTGCAGTCTATTCGCTCT 2765
 851 ProTyrAsnArgAlaValSerLeuAspSerProValSerValGlySerSe 867
 2766 CCATATACCGAGCAGCAGTCTCTGATAGCCCTGTTTCTTGGCTCAAG 2815
 867 rProProValLysAsnIleSerAlaPheProMetLeuProLysGlnProM 884
 2816 TCCTCCAGTAAAAAATATCATGCTCTTCCCATGTTACCAAGCAACCCA 2865
 884 eLeuGlyGlnAsnProArgMetMetLaspSerGlnLysasnTyrgLysSer 900
 2866 TGTGTGGGTGGGAATCCAGAGATGATGATAGTCAGGAAAAATTAATGGCTCA 2915
 901 SerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProSerSe 917
 2916 AGTATGGTGGGCCCAACCGAAATGTGACTGTGACTCAGACTCTCTCTCTC 2965
 917 rGlyAspTrpGlyLeuProAsnSerLysAlaGlyArgMetGluProMetA 934
 2966 AGGAGAGCTGGGGCTTACCAAACTCAAGCGCGGACAGATGGAACCTATGA 3015
 934 snSerAsnSerMetGlyArgProGlyGlyValAspTyrAsnThrSerLeuPro 950
 3016 ATTCAACTCTCATGGAAGACAGAGAGAGATTAATATCTTCTTATCC 3065
 951 ArgProAlaLeuGlyLysSerIleProThrLeuProLeuArgSerAsnSe 967
 3066 AGACCTGCACTGGGTGGCTCTATTCCACACTTGCCTTGGGTCAATAG 3115
 967 rIleProGlyAlaArgProValLeuGlnGlnGlnGlnMetLeuGlnM 984
 3116 CATATCCAGGTGGACAGACAGTATTGGACAGCAGCAGCATAGTCTTCAAA 3165
 984 eArg.....ProGlyGluIleProMetGlyMetGlyAlaAsnPro 997
 3166 TGACTCCCCCAGGGCTGTGGAAATCCCATGGGAATGGGGCTAATCCC 3215
 998 TyrGlyGlnAlaAlaSerAsnGlnLeuGlySerTrpProAspGlyMe 1014
 3216 TATGGCCAAAGCAGCATCTAACAACACTGGTCTCGGCCCATGGCAT 3265
 1014 tLeuSerMetGluGlnValSerHisGlyThrGlnAsnArgProLeuLeuA 1031
 3266 GTTGTCTCATGGAAACAAGTTTCTCATGACACTCAAAATAGGCTCTTCTTA 3315
 1031 rAsnSerLeuAspAspLeuValGlyProProSerAsnLeuGlnGlyGln 1047
 3316 GGAATTCCTCGATGATCTGTGTGGGCACCTTCCAACTGGGAAGGCCAG 3365
 1048 SerAspGluArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnTh 1064
 3366 AGTGACGAAGAAGACATTAATGGACAGCTGCACACTCTTCTCAGCAACAC 3415
 1064 rAspAlaThrGlyLeuGlnGluIleAspArgAlaLeuGlyIleProGluL 1081
 3416 AGATGGCACAGGCTGTGAAGAAATTAAGACAGCTTGGGCATTCCTGAAC 3465
 1081 euValAsnGlnGlyGlnAlaLeuGlnProLysGlnAspAlaPheGlnGly 1097
 3466 TTGTCAATCAGGACAGGCAATTAGAGCCCAACAGGATGCTTCCAAAGGC 3515
 1098 GlnGlnAlaAlaValMetMetAspGlnLysAlaGlyLeuArgGlyGlnTh 1114

3516 CAAGAGCAGCAGCATTAATGATGATCAGAAAGCAGATTAATATGAGACAGC 3565
 1114 rTyProAlaGlnGlyProProMetGlnGlyGlyPheHisLeuGlnGly 1131
 3566 ATACCCAGCAGCAGGGGCTCCCAATGCAGAGAGGCTTTCATCTTCAGGAGC 3615
 1131 InSerProSerPheAsnSerMetMetLaspGlnMetAsnGlnGlnGlyAsn 1147
 3616 AATCACCATCTTTTAACCTATATGATGATGATGAATGAACCGAAGGCAT 3665
 1148 PheProLeuGlnGlyMetHisProArgAlaAsnIleMetArgProArgTh 1164
 3666 TTTCCTCTCCAAAGATGCAACCCAGCAGCAATCATGAGACCCCGGAC 3715
 1164 rAsnThrProLysGlnLeuArgMetGlnLeuGlnGlnArgLeuGlnGly 1181
 3716 AAACACCCCAACCAACTTGAATGACAGTTCAGCAGAGGCTCAGAGGCC 3765
 1181 InGlnPheLeuAsnGlnSerArgGlnAlaLeuGlnLeuLysMetGluAsn 1197
 3766 AGCAGTTTGTGAATCAGAGCCGACAGGCACTTGAATGAATGAATGAATGA 3815
 1198 ProThrAlaGlyGlyAlaAlaValMetArgProMetMetGlnPro..... 1212
 3816 CTTACTGCTGTGTGCTGCTGGGTGATGAGGCTATGATGAGGCCCAAGT 3865
 1213GlnGlnGlyPheLeuAsnAlaGlnMetValAlaGlnArgSera 1227
 3866 GACCTCCACAGCAGGTTTCTTAATGCTCAAAATGGTCGCCCAACGACAGA 3915
 1227 rGlyLeuLeuSerHisHisPheArgGlnGlnArgValAlaMetMetMet 1243
 3916 GAGAGCTGGTAAAGTCAATCACTTCGACACAGAGGGGTATGATGATG 3965
 1244GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1258
 3966 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 4015
 1258 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnAlaPheSerPro 1275
 4016 GCACACAGCAACACAGCAGCAGCAGCAGCAAAACCCAGGCTTACGCCCACT 4065
 1275 rAsnValThrAlaSerProSerMetLaspGlyLeuLeuAlaGlyProThr 1291
 4066 CTATGAGACTGCTTCCCCAGCATGATGAGGCTTTTGGACGACCCACA 4115
 1292 MetProGlnAlaProProGlnGlnPheProTyrGlnProAsnTyrgLyme 1308
 4116 ATGCCAACAGCTCTCCGCAACAGTTTCCATATCAACCAATTAATGGAAT 4165
 1308 tGlyGlnGlnProAspProAlaPheGlyArgValSerSerProProAsnA 1325
 4166 GGGACAAACAACAGATCCAGCTTGTGGTGCAGTGTAGTCCCAATG 4215
 1325 tMetMetSerSerArgMetGlyProSerGlnAsnProMetMetGlnHis 1341
 4216 CAATGATGTCTCAAGAAATGGTCCCTCCAGATCCCATGATGCAACAC 4265
 1342 ProGlnAlaAlaSerIleTyrGlnSerSerGlnMetLysGlyTrpProse 1358
 4266 CCGCAGGCTCATCATTAATCATAGTCTCTGAGAAATGAAGGCTGGGCATC 4315
 1358 rGlyAsnLeuAlaArgAsnSerSerPheSerGlnGlnGlnPheAlaHisG 1375
 4316 AGGAAATTTGGCCAGGACAGCTCTTTTCCCGCAGCAGTGTGGCCACAC 4365
 1375 InGlyAsnProAlaValTyrSerMetValHisMetAsnGlySerSerGly 1391
 4366 AGGGAGATCTCGAGTGTATGATGATGATGATGATGATGATGATGATGATG 4415
 1392 HisMetGlyGlnMetAsnMetAsnProMetProMetSerGlyMetProMe 1408
 4416 CACATGGAGAGATGAACATGAACCCCATGCTGCTGGCATGCTCTAT 4465

1408 tglYProaspGlnIysTyrCys 1415
|||||
4466 GGGTCTGTGATCAGAAATACGTC 4487

seq_name: /cgn2_6/pcdata/1/pna/US094_COMB.seq:us-09-440-612-1

seq_documentation_block:

; Sequence 1, Application US/09440612
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-3 EXPRESSION
; FILE REFERENCE: RTS-0042
; CURRENT APPLICATION NUMBER: US/09/440,612
; CURRENT FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 6754
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (184)..(4422)
US-09-440-612-1

alignment_scores:

Quality: 7234.50 Length: 1427
Ratio: 5.175 Gaps: 3
Percent Similarity: 97.968 Percent Identity: 97.758

alignment_block:

US-09-041-994-2 x US-09-440-612-1 ..

Align seg 1/1 to: US-09-440-612-1 from: 1 to: 6754

1 MetserGlyLeuGlyGluAsnLeuAspProLeuAlaSerAspSerArgLy 17
|||||
184 ATGAGTGATTAGAGAAACTGTGATCCACTGCGCACTGATTCACGAAA 233
17 sArgLysLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGlyG 34
|||||
234 ACGCAAAATGTCATGTGATCTCCAGCAAGCTTACCTGCGAGGTG 283
34 LuLysArgArgArgGluGlnGluSerLysTyrIleGluGluLeuAlaGlu 50
|||||
284 AAAAAGGAGACGGAGCAGGAAAGTAATATATTGAAGAAATGGCTGAG 333
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67
|||||
334 CTGATATCTCGCAATCTTATGATATGTGACAAATTCATGTCAACCCGA 383
67 pLysCysAlaIleLeuLysGluThrValArgGlnIleArgGlnIleLysG 84
|||||
384 TAAATGTGCGATTTTAAAGAAACAGTAAGACAGTACGTCAATAAAG 433
84 LuGlnGlyLysThrIleSerAsnAspAspValGlnLysAlaAspVal 100
|||||
434 AGCAAGGAAAAACTATTTCCAAATGATGATGATTCAAAAGCCGATGTA 483
101 SerSerThrArgLysGlnGlyValIleAspLysAspSerLeuGlyProLeu 117
|||||
484 TCTTCTACAGGGCAGGGAGGATTATGATAAACCTCTTAAGAGACCGCTTT 533
117 uLeuGlnAlaLeuAspGlyPheLeuPheValIValAsnArgGlnAlaAsnI 134
|||||
534 ACTTCAGGCAATGGATGGTTCTCTATTGTGTGATGATGAGACGGAACA 583
134 IValIleValSerGluAsnValIThrGlnTyrLeuGlnTyrLysGlnGlu 150
|||||
584 TTGTATTGTGATCAGAAAAATGTCAACAAATACGTCAATATATAGCAAGAG 633

151 AspLeuValAsnThrSerValTyrAsnIleLeuHISGluAspArgLy 167
|||||
634 GACCTGGTTAAACACAAAGTTTACAAATATCTTACATGAAGAAGACAGAAA 683
167 sAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTyrP 184
|||||
684 GGATTTTCTTAAGAAATTTACCAAAATCTACAGTATATGAGACTTCCGTGA 733
184 hAsnGluProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200
|||||
734 CAAATGACACCCCAAGACAAAAAGCCATCATTTAAATTTGCCGTATGTTG 783
201 MetLysThrProHisAspIleLeuGlnAspIleAsnAlaSerProGluMe 217
|||||
784 ATGAAGAACACACATGATATTTCTGAAGACATTAACGCCACTCTGGAAT 833
217 LArgGlnArgTyrGluThrMetGlnCysPheAlaLeuSerGlnProArg 234
|||||
834 GCGCCAGAGATATGAACAAATGCAAGTGTGCTTGGCTGTCTAGCCACGAG 883
234 IAmMetGluGluGlyGluAspLeuGlnSerCysMetIleCysValAla 250
|||||
884 CTATGATGAGGAAGGGGAAGATTGCAATCTGTGATGATCTGTGTGCA 933
251 ArgArgIleThrThrGlyGluArgThrPheProSerAsnProGluSerPh 267
|||||
934 CGCCGCAATCTACAGAGAAAGAAACATTTCCATCAAACTCGAGAGCTT 983
267 eIleThrArgHisAspLeuSerGlyLysValIValAsnIleAspThrAsn 284
|||||
984 TATTACAGACATGATCTTTAGGAAAGCTTGTCAATATATATCAAAAT 1033
284 eLeuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300
|||||
1034 CACTGAGATCCTCCATGAGCGCTGCTTGAAGATATATACGAAGGTGT 1083
301 IleGlnArgPhePheSerLeuAsnAspGlyGlnSerThrPserGlnLysAr 317
|||||
1084 ATTCAAGATTTTTTATGCTTAATGATGGCAGTCATGTGCCAGAAACG 1133
317 gHisTyrGlnGlu.....AlaTyrL 324
1134 TCACATATAAAGAGTTACCACTGATGGATATTTTCCCAACACACTTATTC 1183
324 euAsnGlyHisAlaGluThrProValTyrArgPheSerLeuAlaAspGly 340
|||||
1184 TTAATGGCCATGACAGAAACCCAGATATGATTCGTTGCGTAGATGA 1233
341 ThrIleValIThrAlaGlnThrLysSerLysLeuPheArgAsnProValTh 357
|||||
1234 ACTATAGTACTGCACAGACAAAAAGCAAACTCTTCGAAATVCCCTGTATAC 1283
357 rAsnAspArgHisGlyPheValSerThrHisPheLeuGlnArgGluAla 374
|||||
1284 AAATGATGACATGGCTTTGTCTCAACCCACTTCTCTCAGAGAAACAGA 1333
374 snGlyTyrArgProAsnProAsnProValGlyGlnGlyIleArgProPro 390
|||||
1334 ATGATATAGACCAAAACCAATCTCTGTTGGACAAAGGATTAAGACCACT 1383
391 MetAlaGlyCysAsnSerSerValGlyLysMetSerMetSerProAsnG 407
|||||
1384 ATGGCTGGATGCAACAGTTTCGTAGCGCGCATGATGATGTGCCAAACCA 1433
407 nGlyLeuGlnMetProSerSerArgAlaTyrGlyLeuAlaAspProserT 424
|||||
1434 AGGCTTAAGATGCCGACAGCAGAGGCTTATGGCTTGACAAACCTTACCA 1483
424 hThrGlyGlnMetSerGlyAlaArgTyrGlyLysSerAsnIleAla 440
|||||
1484 CCACAGGCGACATGATGAGCTAGGTATGGGGTTCCAGTAACATAGACT 1533
441 SerLeuThrProGlyProGlyMetGlnSerProSerSerTyrGlnAsnAs 457

```

1534 TCATTGACCCCTGGCCAGGATGCAATCACCATCTTCTACAGAACAA 1583
457 nasnYrGlyLeuAsnMetSerSerProHisGlySerProGlyLeuA 474
1584 CAACTATAGCTACATGATGAGTCCGCCACATGGAGTCTGCTTGTG 1633
474 lArProsnGlnInsnIleMetIleSerProArgAsnArgGlySerPro 490
1634 CCCCACACGACAGAAATATCATATTTCTCTCTGTAATCGTGGAGTCA 1683
491 LysIleAlaSerHisGlnPheSerProValAlaGlyValHisSerProme 507
1684 AAGATAGCCCTACATCAGTTTCTCTGTTGACAGGTGTGACCTCCAT 1733
507 tAlaSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerLeus 524
1734 GGCATCTTCTGGCAATACCTGGGAACACACCTTTTCCAGCAGCTCTCA 1783
524 eAlaIleuGlnAlaIleSerGlyValGlyThrSerLeuLeuSerThr 540
1784 GTCCCTCGACAGCCATCAGTGAAGGTGTGGGACTTCCCTTTATCTACT 1833
541 LeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleTh 557
1834 CTGTCAATCACAGGCCCAAAATTTGAACTCTCCCAATATGAATATAC 1883
557 rGlnProSerLysValSerAsnGlnAspSerLysSerProLeuGlyPheT 574
1884 CCAACCAAGAAAGTAAGCAATCAGATTCACAGAGTCCCTGCGCTTTT 1933
574 yICysAspGlnAsnProValGluSerSerMetCysGlnSerAsnSerArg 590
1934 ATTGCACCAAAATCCAGTGAAGATTCATGTCTCAATCAATATGAGCA 1983
591 AspHisLeuSerAspLysGlnSerLysGlnSerSerValGluGlyAlaG 607
1984 GATCACCTCAGTGACAAAGAAAGTAAGGAGAGCAGTGTGAGGGGAGCA 2033
607 uAsnGlnAlaGlyProLeuGlnSerLysGlyHisLysLysLeuLeuGln 624
2034 GAATCAAGGGGCTTGTGAAAGCAAGGTCATAAAAAATTAATCTGACGT 2083
624 euleuThrCysSerSerAspAspArgLysHisSerSerLeuThrAsnSer 640
2084 TACTTACCTGTTCTTCTGATGACGGGGTCAITCTCTGACCAACTCC 2133
641 ProLeuAspSerSerCysLysGlnSerSerValSerValThrSerProse 657
2134 CCCCTGATTCAGATTGTAAGAATCTTCTGTAGTGTACACAGCCCTC 2183
657 rGlyValSerSerSerThrSerGlyValSerSerThrSerAsnMetH 674
2184 TGGAGTCTCTCTCTACATCATCTGAGGAGATCTCTACATCCAAATATGC 2233
674 fsgLysSerLeuLeuGlnGlnLysHisArgTleuHisLysLeuLeuGln 690
2234 ATGGGTCACTGTTACAGAGAGACACGGATTTTGCAACAATGCTGCAG 2283
691 AsnGlyAsnSerProAlaGlyValAlaLysIleThrAlaGlnAlaThrG 707
2284 AATGGGAATTCACACAGCTGAGTGAAGCAATTAAGTCAACAACCACTGG 2333
707 yLysAspThrSerSerIleThrSerCysGlyAspGlyAsnValValLysG 724
2334 GAAGAGACACGACAGTATACTCTGTGGGACGGAATGTGTCAAGC 2383
724 lngLlnGlnLeuSerProLysLysLysGlnAsnAsnAlaLeuLeuAlaGlyT 740
2384 AGGAGCAGCTAAATCTCTAAGAAAGAGAGAAATATGACATCTTAGATAC 2433
741 LeuLeuAspArgAspAspProSerAspAlaLeuSerLysGlnLeuGlnPr 757
1534 TCATTGACCCCTGGCCAGGATGCAATCACCATCTTCTACAGAACAA 1583
457 nasnYrGlyLeuAsnMetSerSerProHisGlySerProGlyLeuA 474
1584 CAACTATAGCTACATGATGAGTCCGCCACATGGAGTCTGCTTGTG 1633
474 lArProsnGlnInsnIleMetIleSerProArgAsnArgGlySerPro 490
1634 CCCCACACGACAGAAATATCATATTTCTCTCTGTAATCGTGGAGTCA 1683
491 LysIleAlaSerHisGlnPheSerProValAlaGlyValHisSerProme 507
1684 AAGATAGCCCTACATCAGTTTCTCTGTTGACAGGTGTGACCTCCAT 1733
507 tAlaSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerLeus 524
1734 GGCATCTTCTGGCAATACCTGGGAACACACCTTTTCCAGCAGCTCTCA 1783
524 eAlaIleuGlnAlaIleSerGlyValGlyThrSerLeuLeuSerThr 540
1784 GTCCCTCGACAGCCATCAGTGAAGGTGTGGGACTTCCCTTTATCTACT 1833
541 LeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleTh 557
1834 CTGTCAATCACAGGCCCAAAATTTGAACTCTCCCAATATGAATATAC 1883
557 rGlnProSerLysValSerAsnGlnAspSerLysSerProLeuGlyPheT 574
1884 CCAACCAAGAAAGTAAGCAATCAGATTCACAGAGTCCCTGCGCTTTT 1933
574 yICysAspGlnAsnProValGluSerSerMetCysGlnSerAsnSerArg 590
1934 ATTGCACCAAAATCCAGTGAAGATTCATGTCTCAATCAATATGAGCA 1983
591 AspHisLeuSerAspLysGlnSerLysGlnSerSerValGluGlyAlaG 607
1984 GATCACCTCAGTGACAAAGAAAGTAAGGAGAGCAGTGTGAGGGGAGCA 2033
607 uAsnGlnAlaGlyProLeuGlnSerLysGlyHisLysLysLeuLeuGln 624
2034 GAATCAAGGGGCTTGTGAAAGCAAGGTCATAAAAAATTAATCTGACGT 2083
624 euleuThrCysSerSerAspAspArgLysHisSerSerLeuThrAsnSer 640
2084 TACTTACCTGTTCTTCTGATGACGGGGTCAITCTCTGACCAACTCC 2133
641 ProLeuAspSerSerCysLysGlnSerSerValSerValThrSerProse 657
2134 CCCCTGATTCAGATTGTAAGAATCTTCTGTAGTGTACACAGCCCTC 2183
657 rGlyValSerSerSerThrSerGlyValSerSerThrSerAsnMetH 674
2184 TGGAGTCTCTCTCTACATCATCTGAGGAGATCTCTACATCCAAATATGC 2233
674 fsgLysSerLeuLeuGlnGlnLysHisArgTleuHisLysLeuLeuGln 690
2234 ATGGGTCACTGTTACAGAGAGACACGGATTTTGCAACAATGCTGCAG 2283
691 AsnGlyAsnSerProAlaGlyValAlaLysIleThrAlaGlnAlaThrG 707
2284 AATGGGAATTCACACAGCTGAGTGAAGCAATTAAGTCAACAACCACTGG 2333
707 yLysAspThrSerSerIleThrSerCysGlyAspGlyAsnValValLysG 724
2334 GAAGAGACACGACAGTATACTCTGTGGGACGGAATGTGTCAAGC 2383
724 lngLlnGlnLeuSerProLysLysLysGlnAsnAsnAlaLeuLeuAlaGlyT 740
2384 AGGAGCAGCTAAATCTCTAAGAAAGAGAGAAATATGACATCTTAGATAC 2433
741 LeuLeuAspArgAspAspProSerAspAlaLeuSerLysGlnLeuGlnPr 757
2434 CTGCTGACAGGAGATGATCTAGTATGACATCTCTAAGAACTATACAGCC 2483
757 oGlnValGluGlyValAspAsnLysMetSerGlnCysThrSerSerThr 774
2484 CCAAGTGGAAAGAGTGTATATAAAGAGTCAAGTGCACAGCTCCACCA 2533
774 lProSerSerSerGlnGlnLysAspProLysIleLysThrGlnThrSer 790
2534 TTCTTACTCAAGTCAAGAGAGAAAGACCTTAATAATTAAAGACAGAGACAGT 2583
791 GluGlnGlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLe 807
2584 GAAGAGGAGATCTGAGACTTGGAATATCTAGATGCTATTCTTGATGATCT 2633
807 uThrSerSerAspPheTyrAsnAsnSerIleSerSerAsnGlySerHisL 824
2634 GACTAGTCTGACTTTTACATATATTCATATCTCTCAAAATGATGATCATC 2683
824 euGlyThrLysGlnGlnValPheGlnGlyThrAsnSerLeuGlyLeuLys 840
2684 TGGGAGCTAAGCAACAGCTTTTCAGAGAACTAATTTCTGTGGGTTTGAA 2733
841 SerSerGlnSerValGlnSerIleArgProProTyrAsnArgAlaValSe 857
2734 AGTTCAAGTCTGCGAGTCTATTCGTCCTCCATATACCGACAGTGTCTC 2783
857 rLeuAspSerProValSerValGlySerSerProProValLysAsnIleS 874
2784 TCTGGATAGCCCTGTTCTGTGGCTCAAGTCTCCAGTAAAAAATATCA 2833
874 eAlaPheProMetLeuProLysGlnProMetLeuGlyGlyAsnProArg 890
2834 GTGCTTTTCCCATGTTTACCAAGCAACCCATGTTGGTGGGAATCCAGCA 2883
891 MetMetAspSerGlnGlnAsnYrGlySerSerMetGlyLysProAsnAr 907
2884 ATGATGATATGTACAGAAATTAATGCTCAAGTATG..... 2919
907 gAsnValThrValThrGlnThrProSerSerGlyAspThrGlyLeuProA 924
2920 .....GGAGACTGGGCTTACCA 2938
924 snSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyArg 940
2939 ACTCAAGGCGCGGAGATGAACTATGAATTCATAAATCTCAATGGGAGCA 2988
941 ProGlyLysAspTyrAsnThrSerLeuProArgProAlaLeuGlyLysE 957
2989 CCAAGAGAGATTAATATACTTCTTTTACCAGACCTGCACTGGGTGCTC 3038
957 rIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgProV 974
3039 TATTTCCACATTCCTCTTCTGCTTAATACATACACAGGTGCAGAGACAG 3088
974 aLeuGlnGlnGlnGlnGlnMetLeuGlnMetArgProGlyValIlePro 990
3089 TATTTGACACGACAGACAGATGCTTCAATATGAGCCGTGTGAATATCCC 3138
991 MetGlyMetGlyAlaAsnProTyrGlyGlnAlaAlaIleAsnGlnLe 1007
3139 ATGGGAATGGGGCTAATCCCTATGGCCAGCAAGCACTATACCACT 3188
1007 uGlySerThrProAspGlyMetLeuSerMetGluGlnValSerHisGlyT 1024
3189 GGGTTCCTGCGCCGATGAGCATGTTGTCCATGAGAAACAATTTCTCATGGCA 3238
1024 hrcGlnAsnArgProLeuLeuArgAsnSerLeuAspLeuValGlyPro 1040
3239 CTCAAAATAGCCCTCTTTAGGAATTTCCCTGGATGATCTTGTGGGCCA 3288
1041 ProSerAsnLeuGlnGlyGlnSerAspGluArgAlaLeuLeuAspGlnLe 1057
3289 CTTTCAACTGGAAGCCCAAGTGAAGAAAGACATTAATGACACAGCT 3338

```


1057 uHisThrLeuSerAsnThrAspAlaThrGlyLeuGlnGluIleAspA 1074
 |||||
 3339 GCACACTCTTCTCAGCAACACAGATGCCAGGCTCGAAGAAATTGACA 3388
 |||||
 1074 rGAlaLeuGlyTLeProGluLeuValAsnGlnGlyGlnAlaLeuGluPro 1090
 |||||
 3389 GAGCTTTGGCATTTCTGAACCTGTCAATCAGGAGCAGCATTTAGAGCC 3438
 |||||
 1091 LysGlnAspAlaPheGlnGlyGlnGlnIleAlaValMetMetAspGlnIly 1107
 |||||
 3439 AAACGAGATGCTTTCAGAGCCCAACAGACAGCATTAATGATGATCAGAA 3488
 |||||
 1107 sAlaGlyLeuTyrglyGlnThrTy-ProAlaGlnGlyProPrometGlnG 1124
 |||||
 3489 GCGAGATTAATATGAGACAGACATACCCAGACAGGCGCTCCAAATGACAAG 3538
 |||||
 1124 LysGlyPheHisLeuGlnGlnGlnSerProSerPheAsnSerMetMetAsn 1140
 |||||
 3539 GAGGCTTTCATCTTCAGGAGCATACCATCTTTAACTGTATGATGAT 3588
 |||||
 1141 GlnMetAsnGlnGlnGlyAsnPheProLeuGlnGlyMetHisProArgAl 1157
 |||||
 3589 CAGATGAACACGACAGCAATTTCTCTCCAGAGAAATGCCACCGACGAC 3638
 |||||
 1157 aAsnIleMetArgProArgThrAsnThrProLysGlnLeuArgMetGlnL 1174
 |||||
 3639 CAACTCATGTAGACCCCGACAAACACCCCAACCACTTAAATATGACAGC 3688
 |||||
 1174 euGlnGlnArgLeuGlnGlnGlnGlnPheLeuAsnGlnSerArgGlnAla 1190
 |||||
 3689 TTCAGCAAGGCTGCAGGCGCCAGCGAGTTTGAATCAGACCGCAGACGCA 3738
 |||||
 3739 CTGGAATTGAAATAAGAAACCCCTACTGCTGTGCTGCGCGTGATGAG 3788
 |||||
 1207 gPrometMetGlnProGlnGlnGlnGlyPheLeuAsnAlaGlnMetValAlaG 1224
 |||||
 3789 GCCATATGTAGACGCCCGCAGAGGTTTCTTAATCTCAAAATGCGTGGCC 3838
 |||||
 1224 InArgSerArgGluLeuLeuSerHisHisPheArgGlnGlnArgValAla 1240
 |||||
 3839 AACCGACAGAGAGCTGCTAAGTCACTCACTCCGCAACAGAGGCTGGCT 3888
 |||||
 1241 MetMetMet.....GlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1255
 |||||
 3889 ATGATGATGACAGACAGCAACAGCAGCAGCAGCAGCAGCAGCAGCA 3938
 |||||
 1255 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnAlaPheS 1272
 |||||
 3939 ACAGCAACAGCAACAGCAACAGCAGCAACAGCAACAGCAACAGCAACAGCA 3988
 |||||
 1272 ePProProProAsnValThrAlaSerProSerMetAspGlyLeuLeuAla 1288
 |||||
 3989 GCCACCTCCCTAATGTGACTGCTTCCCAAGATGATGGCTTTGGCA 4038
 |||||
 1289 GlyProThrMetProGlnAlaProProGlnGlnPheProTyrglnProAs 1305
 |||||
 4039 GGACCCCAAAATGCCACAAAGCTCTCCGCAACAGTTCCATATCAACCAAA 4088
 |||||
 1305 nTyrglyMetGlyGlnGlnProAspProAlaPheGlyArgValSerSerP 1322
 |||||
 4089 TTATGGATGGGACAAACAGATCCAGCCTTTGGTGAAGTGTAACTGC 4138
 |||||
 1322 rOProAsnAlaMetMetSerSerArgMetGlyProSerGlnAsnProMet 1338
 |||||
 4139 CTCCCAATGCAATGATGTGCTCAAGAAATGGGCTCCCTCCCAAGATCCCATG 4188
 |||||
 1339 MetGlnHisProGlnAlaAlaSerTLeTyrglnSerSerGluMetLysGln 1355
 |||||
 4189 ATGCAACACCGCAGGCTGCATCCATCTATCAGTCTCAGAAATGAAAGG 4238

1355 yTrpProSerGlyAsnLeuAlaArgAsnSerSerPheSerGlnGlnGlnP 1372
 |||||
 4239 CTGGCATTCAGGAAATTTGGCCAGAACAGCTCTTTTCCCGCAGCAGT 4288
 |||||
 1372 heAlaHisGlnGlnGlnProAlaValTyrSerMetValHisMetAsnGly 1388
 |||||
 4289 TTGGCCACCGAGGGAATCTTCGATGTATGATGTGTGCACATGTAATGGC 4338
 |||||
 1389 SerSerGlyHisMetGlyGlnMetAsnMetAsnProMetPrometSerGln 1405
 |||||
 4339 AGCACTGGGCATCATGAGCAGATGAAACATGAACCCCATGTCCCATGTCTCG 4388
 |||||
 1405 yMetProMetGlyProAspGlnLysTyrgly 1415
 |||||
 4389 CATGCCATATGGGTCTCGATCAGAAATACTGTC 4419

seq_name: /cgn2_6/ptdata/1/pna/US094_COMB.seq:US-09-440-612-3

seq_documentation_block:
 ; Sequence 3, Application US/09440612A
 ; GENERAL INFORMATION:
 ; APPLICANT: C. Frank Bennett
 ; APPLICANT: Lex M. Cowsett
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-3 EXPRESSION
 ; FILE REFERENCE: RTS-0042
 ; CURRENT APPLICATION NUMBER: US/09/440,612A
 ; CURRENT FILING DATE: 1999-11-15
 ; NUMBER OF SEQ ID NOS: 49
 ; SEQ ID NO 3
 ; LENGTH: 6754
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (184)..(4422)
 US-09-440-612-3

alignment_scores:
 Quality: 7234.50 Length: 1427
 Ratio: 5.175 Gaps: 3
 Percent Similarity: 97.968 Percent Identity: 97.758

alignment_block:
US-09-041-994-2 x US-09-440-612-3 ..

Align seg 1/1 to: US-09-440-612-3 from: 1 to: 6754

1 MetSerGlyLeuGlyGluAsnLeuAspProLeuAlaSerAspSerArgly 17
 |||||
 184 ATGAGTGGATTAGAGAAAACTTGGATCCACTGGCCAGTGAATTCACGAAA 233
 |||||
 17 sArgLysLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGlyG 34
 |||||
 234 ACGCAATTTGCCATGTGATACCTCAGACAAAGTCTTACCTGCATGGTG 283
 |||||
 34 LuLysArgArgArgGluGlnGlnSerLysTyrlLeGluGluLeuAlaGlu 50
 |||||
 284 AAAAAGCAGAGACGGGAGCAGGAAAGTAAATATATGAAAGAAATGGCTGAG 333
 |||||
 51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67
 |||||
 334 CTGATATCTGCCAATCTTAGTGAATATGACAAATTTCAATGTCAACACAGA 383
 |||||
 67 PLYSCysAlaIleLeuLysGluThrValArgGlnIleArgGlnIleLysG 84
 |||||
 384 TAAATGTGCGATTTAAAGGAACAGTAAAGACAGTACGTCAATATAAAG 433
 |||||
 84 LuGlnGlyLysThrIleSerAsnAspAspValGlnLysAlaAspVal 100
 |||||
 434 AGCAAGAGAAAACTATTTCCAATGATGATGTTCACAAAAGCCGATGTA 483
 |||||
 101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeuLe 117


```

|||||
484 TCCTCTACAGGCGAGGAGTTATGATAAGACTCTTAGACCGCTTTT 533
117 ULEUGINAlaleuaspolyPheleuphevalValAsnArgLUAlAsnI 134
|||||
534 ACTTCAGGACTTGGATGTTCTTATTTGGTGCAATCGAGACGGAACA 583
134 lValPheValSerLUAsnValThrGlnTyrLeuGlnTyrLysGlnLU 150
|||||
584 TTGTATTGTCATCAGAAATGTCACAAATACCTGCATATATACCAAG 633
151 AspleuValAsnThrSerValTyrAsnIleleuHisGlnLUAspArgLy 167
|||||
634 GACCTGGTTAAACACAGTGTATTACATATCTTACATGAGAGACAGAAA 683
167 sAsPheleuLysAsnLeuProLysSerThrValAsnGlyValSerTrpT 184
|||||
684 GGATTTTCTTAAGAAATTACCAAAATCTACAGTTAATGAGTTCTTGA 733
184 hrAsnGluProGlnArgGlnLysSerHisThrPheAsnGysArgMetLeu 200
|||||
734 CAATGAGACCCAAAGACAAAAGCCATTTAATTTGCGGTATGTTG 783
201 MetLysThrProHisAspIleleuGluAspIleAsnAlaSerProGluMe 217
|||||
784 ATGAAACACACACATGATATTCTGGAACATAAAGCCAGTCTGAAT 833
217 tArgGlnArgTyrGlnThrMetGlnGysPheAlaLeuSerGlnProArg 234
|||||
834 GCGCCACAGATATGAAACAATGCGAGTCTTCCCTCTCAGCCAGCAG 883
234 lMetMetGlnGlnGluAspLeuGlnSerGysMetIleGysValAla 250
|||||
884 CTATGATGAGGAGGAGGAGATTGCAATCTGTATGATCTGTGTGCA 933
251 ArgArgIleThrThrGlyGluArgThrPheProSerAsnProGluSerPh 267
|||||
934 CGCGCATTTACTACAGGAAAGAACATTTCCATCAAAACCTGAGACTT 983
267 eIleThrArgHisAspLeuSerGlyLysValValAsnIleAspThrAsnS 284
|||||
984 TATTTACAGACATGATCTTTCAGGAAAGTTGCAATATATGATCAAT 1033
284 eIleuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300
|||||
1034 CACTGAGATCTCCATGAGCGCTGGCTTTGAAAGTATATATCCGAAGGT 1083
301 IleGlnArgPhePheSerLeuAsnAspGlyGlnSerTrpSerGlnLysAr 317
|||||
1084 ATTCAGAGATTTTACTCTAAATGATGGCAGTCATGCTCCGAAACG 1133
317 gHisTyrGlnLU.....AlaTyrL 324
|||||
1134 TCATATCAAGAAAGTTACAGTATGGATATTTTCCCAACAGCTTATC 1183
324 euAsnGlnHisAlaGluThrProValTyrArgPheSerLeuAlaAspGly 340
|||||
1184 TTAAATGGCCATGCAAGAAACCCAGTATATGATTCGTTGGCTGATGGA 1233
341 ThrIleValThrAlaGlnThrLysSerLysLeuPheArgAsnProValTh 357
|||||
1234 ACTATAGTGACTGCACAGACAAAAAGCAAACTCTCCGAAATCTGTATAC 1283
357 rAsnAspArgHisGlyPheValSerThrHisPheLeuGlnArgGlnLU 374
|||||
1284 AAATGATGACATAGCTTGTCTCAACCCACTTCTCAGAGAAACAGA 1333
374 snGlyTyrArgProAsnProAsnProValGlyGlnGlyIleArgProPro 390
|||||
1334 ATGGATATAGCCAAACCAAAATCTGTGAGCAAGGATTAGACCACT 1383
391 MetAlaGlyCysAsnSerSerValGlyGlyMetSerMetSerProAsnGI 407
|||||

1384 ATGGCTGATGACAAACAGTTCGGTAGCGGCATGATGTGCCAAACA 1433
407 ngLYleuGlnMetProSerSerArgAlaTyrGlyLeuAlaAspProSerT 424
|||||
1434 AGGCTTACAGATGCCGAGACAGAGGCTTATGGCTTGACAGACCTTGA 1483
424 hrThrGlyGlnMetSerGlyAlaArgTyrGlyLysSerSerAsnIleAla 440
|||||
1484 CCACAGGCGCATATGATGAGACTAGTATGGGGCTTCAGTACATAGCT 1533
441 SerLeuThrProGlyProGlyMetGlnSerProSerSerTyrGlnAsnAs 457
|||||
1534 TCATTTGACCCCTGGGCCAGCATGACATCAACATCTTCTACACAGACAA 1583
457 nasTyrGlyLeuAsnMetSerSerProProHisGlySerProGlyLeuA 474
|||||
1584 CAACCTATAGGCTCAACATGAGTAGCCCCCACAAGGAGATCTGTCTTG 1633
474 lArProAsnGlnGlnAsnIleMetIleSerProArGAsnArgGlySerPro 490
|||||
1634 CCCCACACAGACAGATATATCATGATTTCTCTCTTAATCTGGGAGTCCA 1683
491 LysIleAlaSerHisGlnPheSerProValAlaGlyValHisSerProme 507
|||||
1684 AAGATAGCCCTCACATCAGTTTCTCTGTTGACAGGTGACCTCTCCAT 1733
507 tAlaSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerLeuS 524
|||||
1734 GGCATCTTCTGGCAATACCTGGAAACACAGCTTTTCCAGCAGCTCTCA 1783
524 eAlaIleuGlnAlaIleSerGlnGlyValGlyThrSerLeuLeuSerThr 540
|||||
1784 GTGCCCTGCAAGCCATCAGTAAGTAGTGAGGACTTCCCTTTATTTACT 1833
541 LeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleTh 557
|||||
1834 CTGTCACTACAGGCGCCCAATTGGATTAACCTCCCAATATGATATTAC 1883
557 rGlnProSerLysValSerAsnGlnAspSerLysSerProLeuGlyPheT 574
|||||
1884 CCAACCAAGTAAAGTAAGCAATCAGATTCCAAGATCTCTGGGCTTTT 1933
574 yrcysAspGlnAsnProValGlnSerSerMetGysGlnSerAsnSerArg 590
|||||
1934 ATTGCGACCAAAATCCAGTGAGAGTTCAATGTCTCAGTCAAAATAGAGA 1983
591 AspHisLeuSerAspLysGlnSerLysGlnSerSerValGlnGlyAlaGI 607
|||||
1984 GATCACCTCAGTGCACAAAGAAAGTAAGAGACAGCTGTGAGGGGCGAGA 2033
607 uAsnGlnArgGlyProLeuGlnLysSerLysGlyHisLysLysLeuGlnL 624
|||||
2034 GAATCAAAAGGGGCTCTTGGAAACAAAGTCATTAATAAAATTACTCAGT 2083
624 euLeuThrCysSerSerAspAspArgGlyHisSerSerLeuThrAsnSer 640
|||||
2084 TACTTACCTGTTCTTGTATGATACCGGGGCTCATTCCTCTTGACCAACTCC 2133
641 ProLeuAspSerSerCysLysGlnSerSerValSerValThrSerProse 657
|||||
2134 CCCCTAATTTCAATGTTAAGAAATCTTGTGTATGTTCACACACCCCTTC 2183
657 rGlyValSerSerSerThrSerGlyGlyValSerSerThrSerAsnMet 674
|||||
2184 TGGAGTCTCTCTCTCACTACATCTGAGAGAGTATCTCTACATCAATATGC 2233
674 isGlySerLeuLeuGlnGlnLysHisArgTyrIleLeuHisLysLeuGln 690
|||||
2234 ATGGGTACTGTTCACAGAGAGCACCGGATTTTGCACAAATGTCTGCAG 2283
691 AsnGlyAsnSerProAlaGluValAlaLysIleThrAlaGlnAlaThrGI 707
|||||
2284 AATGGAAATTCACACAGCTGAGGTAGCCAAAGATTACTCAGAGACCACTGG 2333

```

```

707 ylysaPThrSerSerIleThrSerCysGlyAspGlyAsnValValIleG 724
|||||
2334 GAAAGACACACAGCATATACCTCTTGCGGAGCGAATACTTGTCMAAC 2383
724 IuGIuGIuLeuSerProLysLysGluAsnAsnAlaLeuLeuAryTyr 740
|||||
2384 AGGACAGCAGTAACTCCTAAGAAAGAAAGAAATATATCACTTCTTACATAC 2433
741 LeuLeuAspArgAspAspProSerAspAlaLeuSerLysGluLeuGlnP 757
|||||
2434 CTGCTGACAGAGGATGATCCTAGTGTATGCACTCTTAAAGAACTACAGCC 2483
757 oGlnValGluGlyValAspAsnLysMetSerGlnCysThrSerSerThr 774
|||||
2484 CCAAGTGGAAAGAGTGGATTAATAAATGAGTCACTGACACAGCTCCACCA 2533
774 LeProSerSerSerGlnGluLysAspProLysIleLysThrGluThrSer 790
|||||
2534 TTCCTAGCTCAAGTCAAGAGAAAGACCTTAAATTAAGACAGACAGCACT 2583
791 GluGluGlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLe 807
|||||
2584 GAAAGAGGATCTGAGACTTGGATTAATCTAGATGCTATCTTGATGATCT 2633
807 uThrSerSerAspPheTyrAsnAsnSerIleSerSerAsnGlySerHisL 824
|||||
2634 GACTGTCTCTGACTTTTACAAATTAATTCATATCTCAAAATGGTAGTCAATC 2683
824 euGIyThrLysGlnGlnValPheGlnGlyThrAsnSerLeuGlyLeuLys 840
|||||
2684 TGGGGCATAGCAAGCAAGGTGTTCAAGAAACTAATTCCTGGGTTTGAAA 2733
841 SerSerGlnSerValGlnSerIleArgProProTyrAsnArgAlaValSe 857
|||||
2734 ACTTCACACTGTGTGAGTCTATCTCTCCATATACCGAGCACTGTC 2783
857 rLeuAspSerProValSerValGlySerSerProProValLysAsnIles 874
|||||
2784 TCTGGATAGCCCTGTTCTGTGGCTCAAGTCTCCAGTAAATAATATACA 2833
874 exAlaPheProMetLeuProLysGlnProMetLeuGlyLysAsnProArg 890
|||||
2834 GTGCTTCCCATGTTTACCAAGAAACCATGTTGGGTGGGAATCCACAGA 2883
891 MetMetAspSerGlnGluAsnTyrGlySerSerMetGlyGlyProAsnAr 907
|||||
2884 ATGATGGATAGTCAGGAAATATGCTCAAGTATG..... 2919
907 gasnValThrValThrGlnThrProSerSerGlyAspTrpGlyLeuProA 924
|||||
2920 .....GGAGACTGGGCTTAACCA 2938
924 snSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyArg 940
|||||
2939 ACTCAAAAGCCGCGAAGATGAAACGTATGAATTCAAACCTCATGGGAACA 2988
941 ProGlyGlyAspTyrAsnThrSerLeuProArgProAlaLeuGlyGlySe 957
|||||
2989 CCAGAGAGGATTATATCTTCTTACCACAGACCTGCACCTGGGTGGCTC 3038
957 rIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgProV 974
|||||
3039 TATTTCCACATTTGCTTCTGGTCTAATAGCATACCAAGGTGCGAGCCAG 3088
974 alLeuGlnGlnGlnGlnMetLeuGlnMetArgProGlyGluIlePro 990
|||||
3089 TATTCACAACAGACAGCAGATGCTTCAATATGAGCGCTGCTGAATCCCC 3138
991 MetGlyMetGlyAlaAsnProTyrGlyGlnAlaAlaIleSerAsnGlnLe 1007
|||||
3139 ATGGAAATGGGGCTAATCCCTATGGCCAAAGCAGCAGCATTAACCAACT 3188

```

```

1007 uGIySerTrpProAspGlyMetLeuSerMetGluGlnValSerHisGlyT 1024
|||||
3189 GGGTTCTCTGGCCCGATGAGCATGTTGTCCATGTGACAAAGTTTCTCATGGCA 3238
1024 hrcGlnAsnArgProLeuLeuAryAsnSerLeuAspAspLeuValGlyPro 1040
|||||
3239 CTCAAATATAGCCCTCTTCTTAGGAATTCCTCGATGATCTTTGTTGGGCCA 3288
1041 ProSerAsnLeuGluGlyGlnSerAspGluArgAlaLeuLeuAspGlnLe 1057
|||||
3289 CCTTCCAACTCGAAGAGCCAGAGTGCAGAAAGCATTTATGAGCAGACT 3338
1057 uHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuGlnGluIleAspA 1074
|||||
3339 GCACACTTCTTCTCAGCACACAGATGCCACAGCCCTGGAAAGAAATTGACA 3388
1074 rGAlaLeuGlyIleProGluLeuValAsnGlnGlyGlnAlaLeuGluPro 1090
|||||
3389 GAGCTTGGGCGATTCTGTAACCTTGTCAATCAGGAGCAGCATTAGAGCCC 3438
1091 LysGlnAspAlaPheGlnGlnGlnGlnAlaAlaValMetMetAspGlnLy 1107
|||||
3439 AAACAGAGATGCTTTCCAAGGCCAAGAACAGCAGATATGATGATCAGAA 3488
1107 sAlaGlyLeuTyrGlyGlnThrTyrProAlaGlnGlyProProMetGlnG 1124
|||||
3489 GCGAGATTATATGACACACATACCCACACAGGGGCTCCAAATGTCAG 3538
1124 LyGIyPheHisLeuGlnGlnGlnSerProSerPheAsnSerMetMetAsn 1140
|||||
3539 GAGGCTTTCATCTTCAGGACATACCATCTTTTAACTCATGATGATGAT 3588
1141 GlnMetAsnGlnGlnGlnLysAsnPheProLeuGlnGlyMetHisProArgAl 1157
|||||
3589 CAGATGAACACAGCAAGGCATTTTCTCTCCAAAGAAATCCACCCAGAAC 3638
1157 aAsnIleMetArgProArgThrAsnThrProLysGlnLeuArgMetGlnL 1174
|||||
3639 CAACATCATGAGACCCCGGACAAACACCCCAAGCACTTAATAATGCAC 3688
1174 euGlnGlnArgLeuGlnGlnGlnPheLeuAsnGlnSerArgGlnAla 1190
|||||
3689 TTCAGCAGAGGCTGCGAGGCCAGAGTTTGTGATCAGAGCCGACAGGCA 3738
1191 LeuGluLeuLysMetGluAsnProThrAlaGlyGlyAlaAlaValMetAr 1207
|||||
3739 CTTGAATTGAAATGGAATAACCTACTGCTGGTGGCTGCGGTGTGAG 3788
1207 gProMetMetGlnProGlnGlnGlyPheLeuAsnAlaGlnMetValAlaG 1224
|||||
3789 GCCCTATGATGCAGCCCGACAGGAGTTTCTTAAATGCTCAAAATGTGCGCC 3838
1224 InArgSerArgGluLeuLeuSerHisHisPheArgGlnGlnArgValAla 1240
|||||
3839 AACGCGACAGAGAGCGCTAATGATCATCACTTCCGACAAACAGAGGCGCT 3888
1241 MetMetMet.....GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1255
|||||
3889 ATGATGATGCACAGCAGACAGCAACAGCAGCAGCAGCAGCAGCAGCAGCA 3938
1255 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1272
|||||
3939 ACAGCAACAGCAACAGCAACAGCAGCAACAGCAACAGCAACAGCCAGGCTTCA 3988
1272 exProProProAsnValThrAlaSerProSerMetAspGlyLeuLeuAla 1288
|||||
3989 GCCCACTCTCTAATGATGATGCTTCTCCCGACATGATGATGCTTTTGCGA 4038
1289 GlyProThrMetProGlnAlaProProGlnGlnPheProTyrGlnProAs 1305
|||||
4039 GGACCCCAATATGCCAAAGCTCTCCGCAACAGTTTCCATATCAACCAAA 4088
1305 nTyrGlyMetGlyGlnGlnProAspProAlaPheGlyArgValSerSerP 1322

```

```
|||||
4089 TTATGGATGGACACACAGATCCAGCTTGGTGAGTGTCTACTC 4138
1322 roProAsnAlaMetSerSerArgMetGlyProSerGlnAsnPromet 1338
|||||
4139 CTCCCAATGCAATGATGCTGCAAGAAATGGCTCCCTCCAGAAATCCCATG 4188
1339 MetGlnHisProGlnAlaAlaSerIleTyrGlnSerSerGluMetLysGln 1355
|||||
4189 ATGCACACCCCGACAGCTGCATCCATCTATCAAGTCTCAGAAATGAAGGG 4238
1355 yTrpProSerGlyAsnLeuAlaArgAsnSerSerPheSerGlnGlnLys 1372
|||||
4239 CTGGCCATCAGAGAAATTTGGCCAGGAAACAGCTCTTTCCACAGCAGAGT 4288
1372 heaAlaHisGlnGlyAsnProAlaValTyrSerMetValHisMetAsnGly 1388
|||||
4289 TTGGCCACACAGGGAATCTCAGTGTATAGTATAGTGCACATGATGC 4338
1389 SerSerGlyHisMetGlyGlnMetAsnMetAsnPrometPrometSerGln 1405
|||||
4339 AGCAGTGTCTCAGATGGACAGATGAACATGAACCCCATGCTCATGTCTGG 4388
1405 yMetProMetGlyProAspGlnLysTyrCys 1415
|||||
4389 CATGCTATGGGTCTGATCAGAAATACTGC 4419
```

seq_name: /cgn2.6/pdata/1/pna/US096C_COMB.seq:US-09-652-123-8947

seq_documentation_block:

```
; Sequence 8947, Application US/09652123
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600,1186-001
; CURRENT APPLICATION NUMBER: US/09/652,123
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,135
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 9796
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8947
; LENGTH: 7888
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-123-8947
```

alignment_scores:

```
Quality: 7233.50      Length: 1429
Ratio: 5.174          Gaps: 3
Percent Similarity: 97.831      Percent Identity: 97.621
```

alignment_block:

US-09-041-994-2 x US-09-652-123-8947

Align seg 1/1 to: US-09-652-123-8947 from: 1 to: 7888

```
1 MetSerGlyLeuGlyGluAsnLeuAspProLeuAlaSerAspSerArgly 17
|||||
182 ATGAGTGTGATTAAGGAAAACTTGATTCACATGGCAGTATTCACGAAA 231
17 SarGlyLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGlyG 34
|||||
232 ACGCAAAATTCATGCTGATCTCCAGGACGAGGCTTACTCTCAGTGTG 281
34 LuYsArgArgArgGlnGlnGlnSerLysTyrIleGlnGlnLeuAlaLys 50
|||||
282 AAAAAGGAGACGGGACGAGAAAGTAATATATGAAGAAATGGCTGAG 331
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67
|||||
```

```
332 CTGATATCTGCCAAATCTTAGTATATGACATTTCAATATGCAAAACGAG 381
67 pLyScysAlaIleLeuLysGluThrValArgGlnIleArgGlnIleLysG 84
|||||
382 TAAATGTGCGATTTTAAAGAAACGTAAGACAGATAGCTCAAAATAAAG 431
84 LuGlnGlySerThrIleSerAsnAspAspValGlnLysAlaLysVal 100
|||||
432 AGCAAGAAAAACTTTTCCAAATGATGATGATGTTCCAAAAAGCCGATGA 481
101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeu 117
|||||
482 TCTTCTACAGGCGAGGAGTATTGATTAAGACTTCCTTAGAGACCGCTT 531
117 uLeuGlnAlaLeuAspGlyPheLeuPheValValAsnArgGlnAlaAsn 134
|||||
532 ACTTCAGGCATTTGGATGTTCTTAATTTGTGTGATTCAGAACGCAAGAA 581
134 leValPheValSerGluAsnValThrGlnTyrLeuGlnTyrLysGlnGlu 150
|||||
582 TTGATTTTGTATCAGAAATGTCAACAAATACCTGCATATATATAGCAAG 631
151 AspLeuValAsnThrSerValTyrAsnIleLeuHisGlnGlnAspArgly 167
|||||
632 GACCTGGTTAAACACAGGTTTACAAATCTTACATGAAGAAGACAGAAA 681
167 sAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrp 184
|||||
682 GGATTTTCTTAAGAAATTTACCAAAATCTACAGTTAATGAGCTTCTCGA 731
184 hAsnGluProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200
|||||
732 CAATGAGACCCCAAGACAAAGAACCCCTACATTTAATTTGCCGATGTG 781
201 MetLysThrProHisAspIleLeuGlnAspIleAsnAlaSerProGlu 217
|||||
782 ATGAAACACCAACATGATATTTCTGSAAGACATTAAGCCGACATCTCGA 831
217 TarGlnArgTyrGluThrMetGlnCysPheAlaLeuSerGlnProArg 234
|||||
832 GCGCAGAGATATGAAACAAATGACAGTGTGCTGCTCAGCCACGACGAG 881
234 lAmMetGlnGlnGlyGluAspLeuGlnSerCysMetIleCysValAla 250
|||||
882 CTATGATGAGAGAGGAGGAGATTTGCATCTTGTATATCTGTGTGCA 931
251 ArgArgIleThrThrGlyGluArgThrPheProSerAsnProGluSerPh 267
|||||
932 CGCGCATTTACTACAGAGAAACATTTCCATCAAAACCTGAGAGCTT 981
267 eIleThrArgHisAspLeuSerGlyLysValValAsnIleAspThrAsn 284
|||||
982 TATTACCAAGACATGATCTTTTCAGGAAAGTGTCTCAATATACAAATT 1031
284 erLeuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300
|||||
1032 CACTGAGATCTCCATGAGGCTGCTTGAAGATATATATCCGAGGTGT 1081
301 lIleGlnArgPhePheSerLeuAsnAspGlyGlnSerThrPserGlnLys 317
|||||
1082 ATTCAGAGATTTTATGCTAAATGATGGCAGTCATGTGCCAGAAAGG 1131
317 gHisTyrGlnGlu.....AlaTyrL 324
|||||
1132 TCACATATCAAGAAAGTTACCAAGTATGATGATATTTTCCCAACAGCTAT 1181
324 euAsnGlyHisAlaGlnThrProValTyrArgPheSerLeuAlaAspGly 340
|||||
1182 TTAATGCCATGCAAGAAACCCAGTATATGCAATTCCTTGGCTGATGGA 1231
341 ThrIleValThrAlaGlnThrLysSerLysLeuPheArgAsnProValTh 357
|||||
1232 ACTATAGTACTGCACAGCAAAAGCAAACTCTTCCGAAATCTGTATAC 1281
```

357 rAsnAspArgHisGlyPheValSerThrHisPheLeuGlnArgGlyGln 374
 |||||
 1282 AAATGATTCACATGGCTTGTCTCAACCCACTCTTCAGAGAGAACAGA 1331
 374 sngLyThrProAnProAnProValGlyGlnGlyLeuArgProPro 390
 |||||
 1332 ATGGATATAGCCAAACCCAAATCTGTGGACAAAGGATTAAGACCCT 1381
 391 MetAlaGlyCysAsnSerSerValGlyGlyMetSerMetSerProAnG 407
 |||||
 1382 ATGGCGTGCATGCACAGCTTCGGTAGGGCGCATGATGTCCGCAACCA 1431
 407 nGlyLeuGlnMetProSerSerArgAlaArgGlyLeuAlaAspProSe 424
 |||||
 1432 AGGCTTACAGATGCCGACGACAGGGCCCTATGGCTTGCGACACCTA 1481
 424 hThrGlyGlnMetSerGlyAlaArgTyGlyGlySerSerHisLea 440
 |||||
 1482 CCACAGGGCAGATGAGTGAAGTATGGGGTTCAGTAACATAGCT 1531
 441 SerLeuThrProGlyProGlyMetGlnSerProSerSerTyGlnAsn 457
 |||||
 1532 TCATTGACCCCTGGGGCGACGATGCATCCACTTCTTACACAGAACAA 1581
 457 nAsnTyArgGlyLeuAsnMetSerSerProProHisGlySerProGlyLeu 474
 |||||
 1582 CAACATAGCGCTCAACATGAGTAGCCCCACATGGAGTCTCGTCTTG 1631
 474 lAProAnGlnGlnAsnIleMetIleSerProArgAsnArgGlySerPro 490
 |||||
 1632 CCCCAACACGACGAAATATCATGATTTCTCTCGTATATCGTGGAGTCCA 1681
 491 LysIleAlaSerHisGlnPheSerProValAlaGlyValHisSerProme 507
 |||||
 1682 AAGATAGCCTCAATCATGTTTCTCTGTTGGAGGTGCACATCTCCAT 1731
 507 tAlaSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerSerLeus 524
 |||||
 1732 GGCATCTTGTGCAAATACGCGAACACACAGCTTTTCACAGACTCTCTCA 1781
 524 eAlaIleuGlnAlaIleSerGlnGlyValGlyThrSerLeuSerThr 540
 |||||
 1782 GTGCCCTGCAAGCATCATGAGGTGGGACTTCCCTTATTTACT 1831
 541 LeuSerSerProGlyProLysLeuAspAsnSerProAnMetAsnIleTh 557
 |||||
 1832 CTGTATCATCAGGGCCCAATTTGGTAACTCTCCCAATATGATTTAC 1881
 557 tGlnProSerLysValSerAsnGlnAspSerLysSerProLeuGlyPhe 574
 |||||
 1882 CCACCCAAAGTAAAGTAAGCAATCAGGATTCCAAGAGTCTCTGGGCTTT 1931
 574 yrcysAspGlnAsnProValGlnSerSerMetCysGlnSerAsnSerArg 590
 |||||
 1932 ATTGGACCAAAATCCAGTGAAGAGTTCATATGTCAGTCAAAATACCA 1981
 591 AspHisIleLeuSerAspLysGlnSerLysGlnSerSerValGlyGlyAla 607
 |||||
 1982 GATCACCTCAGTACAAAGAAAGTAAGAGACAGCTTGAGGGGGCACA 2031
 607 nAsnGlnAlaGlyProLeuGlnSerLysGlyHisLysLysLeuLeuGln 624
 |||||
 2032 GAATCAAAAGGGTCTTGTGAAAGCAAAAGTCATAAAATTAATCTCAG 2081
 624 euLeuThrCysSerSerAspAspArgGlyHisSerSerLeuThrAsnSer 640
 |||||
 2082 TACTTAACCTGTTCTTCTGATGACGGGGTCATTTCTCTTGACCAATC 2131
 641 ProLeuAspSerSerCysLysGlnSerSerValSerValThrSerProSe 657
 |||||
 2132 CCCCTAGATTCAAGTTGTAAAGATCTTCTGTAGTGCACCAAGCCCTTC 2181

657 rGlyValSerSerSerThrSerGlyValSerSerThrSerAsnMetH 674
 |||||
 2182 TGGAGTCTCTCTCTCTACATCTGGAGAGATATCTTACATCCAAATATC 2231
 674 lsgLySerLeuGlnGlnLysHisArgIleLeuHisLysLeuGln 690
 |||||
 2232 ATGGGTCACTGTTACAGAGAAAGCACCGGATTTTGCACAAAGTGTGCAG 2281
 691 AsnGlyAsnSerProAlaGlyValAlaLysIleThrAlaGlnAlaThrG 707
 |||||
 2282 AATGGGAATTCACACCACTGAGGTAGCCAAAGATTACGCAAGAGCCACTG 2331
 707 yLysAspThrSerSerIleThrSerCysGlyAspGlyAsnValValLysG 724
 |||||
 2332 GAAAGACACACAGCAGTAACTTCTGTGGGGCGGAAATGTTGTTCACAC 2381
 724 lngLglnLeuSerProLysLysGlnAsnAsnAlaLeuLeuArgTy 740
 |||||
 2382 AGGAGCAGCTAAGTCTTAAGAAAGAGAGAAATATGCACTTCTTAGATAC 2431
 741 LeuLeuAspArgAspAspProSerAspAlaLeuSerLysGlyLeuGlnPr 757
 |||||
 2432 CTGTGTGACAGGAGATGATCTAGTATGCACATCTCTAAAGAACTACAGCC 2481
 757 oGlnValGlyGlyValAspAsnLysMetSerGlnCysThrSerSerThr 774
 |||||
 2482 CCAGTGAAGAGAGTGATTAATAATGAGTCAAGTCCACACCTCCACCA 2531
 774 lProSerSerSerGlnGlnLysAspProLysIleLysThrGlnThrSer 790
 |||||
 2532 TTCTTAGCTCAAGTCAAGGAAAGACCTTAATAATTATAGCAGACAGACT 2581
 791 GlnGlnGlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLe 807
 |||||
 2582 GAAAGAGGATCTGAGACCTTGATATCTAGATGCTATCTTGGTGATCT 2631
 807 uThrSerSerAspPheTyAsnAsnSerIleSerSerAsnGlySerHisL 824
 |||||
 2632 GACTACTTCTGACTTTTACAAATAATTCATATCCCAATGAGTACATC 2681
 824 euGlyThrLysGlnGlnValIlePheGlnGlyThrAsnSerLeuGlyLys 840
 |||||
 2682 TGGGACTTAAGCAACAGGTGTTCAAGAACTAATTTCTGGGTTTGAA 2731
 841 SerSerGlnSerValGlnSerIleArgProProTyAsnArgAlaValSe 857
 |||||
 2732 AGTTCACTGCTGTGACGCTATTCCTCTCCATATACGAGAGACTGTC 2781
 857 rLeuAspSerProValSerValGlySerSerProProValLysAsnIleS 874
 |||||
 2782 TCTGATATACCCCTGTTTCTGTGGCTCAAGTCTCCAGTAAATAATATCA 2831
 874 eAlaIlePheProMetLeuProLysGlnProMetLeuGlyLysAsnProArg 890
 |||||
 2832 GTGCTTCCCAATGTTATCCAAAGCAACCCATGTGGGTGGGATCCAA 2881
 891 MetMetAspSerGlnGlnAsnTyArgLysSerSerMetGlyGlyProAsnAr 907
 |||||
 2882 ATGATGATAGTACAGAAATAATATGCTCAAGTATG 2917
 907 gAsnValIleThrValIleGlnIleThrProSerSerGlyAspTrpGlyLeuPro 924
 |||||
 2918GGAGACTGGGCTTACCAA 2936
 924 snSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyArg 940
 |||||
 2937 ACTCAAAAGCCGGCAGAAATGAAACCTATGAATTCAAATCCATGGGAAGA 2986
 941 ProGlyGlyAspTyAsnThrSerLeuProArgProAlaLeuGlyGlySe 957
 |||||
 2987 CCAGAGAGGATTAATATATCTTATACCCACACTGCACCTGGGTGGCTC 3036
 957 rIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgPro 974

17 sarGlyLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGlyG 34
 |||||
 232 ACGCAAAATGGCATGTGATACCTCAGACAGAGTCTTACCTGCAGTGG 281
 34 LuLyArgArgArgGluGlnGlnSerLysTyrIleGluLueuAlaGlu 50
 |||||
 282 AAAAAGCGAGCGGAGCAGAAAGTAATATATGAAAGAAATGGCTGAG 331
 51 LeuIleSerAlaSerLeuSerAspIleAspAsnPheAsnValLysProAs 67
 |||||
 332 CTGATATCTGCCAATCTTAGTGATATGTGCAATTTCAATGTCAAAACCA 381
 67 pLySCysAlaIleLeuLysGluThrValArgGlnIleArgGlnIleLysG 84
 |||||
 382 TAAATGTGGCATTTTAAAGCAACAGTAAGACAGATACGTCAAAATAAAG 431
 84 LuGlnGlyLysThrIleSerAsnAspAspValGlnLysAlaAspVal 100
 |||||
 432 ACCAAGGAAAAACTATTTCCAAATGATGATGATGTTCAAAAAGCCGATGTA 481
 101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeu 117
 |||||
 482 TCTTTACAGGGCAGGAGTATATGATAAAGACTCTTAGACCCGCTTTT 531
 117 uLeuGlnAlaLeuAspGlyPheLeuPheValValAsnArgGlnAlaAsnI 134
 |||||
 532 ACTTCAGGCATTTGATGGTTCCTATTGTGGTGAATCGAGACGAAACA 581
 134 IeValPheValSerGluAsnValThrGlnTyrLeuGlnTyrLysGlnGlu 150
 |||||
 582 TTGTATTGTATCAGAAATGTACACCAATACCTGCATATATAAGCAAGG 631
 151 AspLeuValAsnThrSerValTyrAsnIleLeuHisGlnGluAspArgLy 167
 |||||
 632 GACCTGGTTAACACAAAGTCTTACATATCTTACATGAAGAAAGACGAAA 681
 167 sasPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrpT 184
 |||||
 682 GGATTTTCTTAAGAAATTTACCAAAATCTACAGTTATGAGATTCTCTGA 731
 184 hAsnGluProGlnArgGlnLysSerHisThrPheAsnGlySarGlyMetLeu 200
 |||||
 732 CAAATGAGACCCCAAGACAAAAGCCATACATTTAATGGCCGTATGGTG 781
 201 MetLysThrProHisAspIleLeuGlnAspIleAsnAlaSerProGlu 217
 |||||
 782 ATGAAACACCCACATGATATTTCTGGAAGACATTAACGCCAGTCTGAAT 831
 217 tarGlnArgTyrGluThrMetGlnCysPheAlaLeuSerGlnProArg 234
 |||||
 832 GCGCCAGAGATATGAACCAATGCAAGCTTTGCCCTGCTCAGCCAGAG 881
 234 IAsnMetGlnGluGlyLysAspLeuGlnSerCysMetIleCysValAla 250
 |||||
 882 CTATGATGGAGGAGGGAAGATTTGCAATCTGTATGATCTGTGGGA 931
 251 ArgArgIleThrThrGlyLysArgThrPheProSerAsnProGluSerP 267
 |||||
 932 CGCCCATTAACACAGGAGAAAGACATTTCCATCAACCCCTGAGAGCTT 981
 267 eIleThrArgHisAspLeuSerGlyLysValValAsnIleAspThrAsn 284
 |||||
 982 TATTTACAGACATGATCTTTACAGAAAGCTGTCAATATATAGATACAAAT 1031
 284 eLeuAspSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300
 |||||
 1032 CACTAGATCCCTCCATGAGGCTGCTGTGAAGATATAATCCGAAGCTT 1081
 301 IleGlnArgPhePheSerLeuAsnAspGlyGlnSerTrpSerGlnLysAr 317
 |||||
 1082 ATTCAAGAGATTTTATAGTCTAATATGATGGCAGTATGTGCCAGAAACG 1131

317 gHisTyrGlnGlu.....AlaTyrL 324
 |||||
 1132 TCACATATCAAGAAGTTACCAAGTATGATGATATTTTCCCAACAGCTTATC 1181
 324 euAsnGlyHisAlaGluThrProValTyrArgPheSerLeuAlaAspGly 340
 |||||
 1182 TTAATGGCCATGCAGAAACCCAGTATATGATTTCTGTTGGCTGATGGA 1231
 341 ThrIleValThrAlaGlnThrLysSerLysLeuPheArgAsnProValTh 357
 |||||
 1232 ACTATGTGACGTGCACAGCAAAAAAGCAAACTCTCCGAAATCCTGTAA 1281
 357 rAsnAspArgHisGlyPheValSerThrHisPheLeuGlnArgGluGln 374
 |||||
 1282 AAATGATCAGACATGGCTTGTCTCAACCCACTTCTTGAGAGAGAAACAG 1331
 374 sngLyTyrArgProAsnProAsnProValGlyGlnGlyLysIleArgProP 390
 |||||
 1332 ATGGATATGACCAAAACCCAAATCTGTTGGACAAAGGATATGACACCT 1381
 391 MetAlaGlyCysAsnSerSerValGlyGlyMetSerMetSerProAsnG 407
 |||||
 1382 ATGGCTGGATGCCACAGTTCGGTAGCGGCATGATGATGCCCAAAACCA 1431
 407 nGlyLeuGlnMetProSerSerArgAlaTyrGlyLeuAlaAspProSerT 424
 |||||
 1432 AGCCTTACAGATGCCAGACAGGCGCTATGGCTTGACAGACCTTAGCA 1481
 424 hThrGlyGlnMetSerGlyAlaArgTyrGlyGlySerSerAsnIleAla 440
 |||||
 1482 CCACAGGCGAGATGATGGAGTATGGATGGGGGTTCCAGTAACATAGCT 1531
 441 SerLeuThrProGlyProGlyMetGlnSerProSerSerTyrGlnAsnAs 457
 |||||
 1532 TCATTATACCCCTGGCGCAGGATGCATATCACCATCTCTTACCAGAACCA 1581
 457 nAsnTyrGlyLeuAsnMetSerSerProProHisGlySerProGlyLeu 474
 |||||
 1582 CAACATATAGGCTCAACATAGTAGGCCCCACATAGGAGATGCTGGCTTG 1631
 474 IAspAsnGlnGlnAsnIleMetIleSerProArgAsnArgGlySerPro 490
 |||||
 1632 CCCCAACCCAGCAAGATATCATGATTTCTCTCGTATATGTTGGAGTCCA 1681
 491 LysIleAlaSerHisGlnPheSerProValAlaGlyValHisSerProMe 507
 |||||
 1682 AAGATATGCTCATCATCAGTTTCTCTGTTGCAGAGTGTCACTCCCT 1731
 507 tAlaSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerSerLeus 524
 |||||
 1732 GGCATCTTCTGGCAATACTGGGAACCCAGCTTTTCCACACACCTCTTCA 1781
 524 eAlaLeuGlnAlaIleSerGluGlyValGlyThrSerLeuLeuSerThr 540
 |||||
 1782 GTGCCCTGCAGCCATCAGTGAAGGTGTGGGACTTCCCTTTATCTACT 1831
 541 LeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleTh 557
 |||||
 1832 CTGTATATCCAGGCGCCCAATTTGATTACTCTCCCAATATATAATATTC 1881
 557 rGlnProSerLysValSerAsnGlnAspSerLysSerProLeuGlyPheT 574
 |||||
 1882 CCACCAAGTAAAGTAAGCAATCAGGATTCCAAGAGTCTCTGGGCTTTT 1931
 574 yrcYaspgLInAsnProValGlySerSerMetCysGlnSerAsnSerArg 590
 |||||
 1932 ATTGGACCAAAATATCAGTGGAGAGTTCAATGTGTCAAGTCAATATGACA 1981
 591 AspHisIleuSerAspLysGluSerLysGluSerSerValGlyAlaGly 607
 |||||
 1982 GATCACTCAGTGCAGAAAGAAAGTAAGAGACACAGTGTGAGAGGGGCGAG 2031
 607 uAsnGlnArgGlyProLeuGluSerLysGlyHisLysLysLeuLeuGlnL 624

```

|||||
2032 GAATCAAGGGGCTCTTGGAAAGCAAGGTCATATAAAATTTACTGCAGT 2081
|||||
624 euLeuThrCysSerSerAspAspArgLysSerSerLeuThrAsnSer 640
|||||
2082 TACTTACTCTTCTTCTGATGACCGGGGTCATTCCTCTGACCAACTCC 2131
|||||
641 ProLeuAspSerSerCysLysGlnSerSerValSerValThrSerProse 657
|||||
2132 CCCCTAGATTCAGTTGTAAGAAATCTCTGTATGTGCACACGCCCTC 2181
|||||
657 rGlyValSerSerSerThrSerGlyValSerSerThrSerAsnMetH 674
|||||
2182 TGGAGCTCTCCCTCTACATCTGGAGGAGTACTCTACATCAATATGTC 2231
|||||
674 tsGlySerLeuLeuGlnGlnLysHisArgLysLeuHisLysLeuLeuGln 690
|||||
2232 ATGGGTCACTGTTACAAGAGAACCGGATTTTGCACAAAGTTGCTCAG 2281
|||||
691 AsnGlyAsnSerProAlaGlnValAlaLysThrAlaGlnAlaThrGln 707
|||||
2282 AATGGGAATTCACCAAGCTGAGGTAGCCAAAGATTAAGTGCAGAACGCTGG 2331
|||||
707 LysAspThrSerSerThrSerCysGlyAspGlyAsnValValLysG 724
|||||
2332 GAAAGACACACAGCAGTAACTTCTTGGGGAGGAAATGTTCTCAAGC 2381
|||||
724 lngLglnLeuSerProLysLysLysGlnAsnAlaLeuLeuArgTyr 740
|||||
2382 AGGAGCAGCTAAAGTCTTAAGAAAGAGAAATATGCACTTCTTAAGATAC 2431
|||||
741 LeuLeuAspArgAspAspProSerSerAspAlaLeuSerLysGlnLeuGlnPro 757
|||||
2432 CTGCTGACACAGGATGATCCTAGTATGACACTCTCTAAAGACATACGCC 2481
|||||
757 GlnValGlnGlyValAspAsnLysMetSerGlnCysThrSerSerThrI 774
|||||
2482 CCAAGTGGAAAGAGTGGATATAATAATGAGTCACTGCACGCCAGCTCCACCA 2531
|||||
774 leProSerSerSerGlnGlnLysAspProLysIleLysThrGlnThrSer 790
|||||
2532 TTCTACTACTCAAGTCAAGAGAAAGACCTTAATAATTAAGACAGACAGT 2581
|||||
791 GlnGlnGlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlnLysPle 807
|||||
2582 GAAGAGGGATCTGGAGACTTGGATATCTAGATCTTATTTCTGGTGATCT 2631
|||||
807 urThrSerSerAspPheTyrAsnAsnSerIleSerSerAsnGlySerHisL 824
|||||
2632 GACTAGTTTCGACTTTTACATATATTCATATCTCTCAAAATGTTAGTCATC 2681
|||||
824 euGlyThrLysGlnGlnValPheGlnGlnLysThrAsnSerLeuGlyLeuLys 840
|||||
2682 TGGGGACTAAGCAACAGGTGTTTCAAGAACTAATTTCTTGGGTTTGAA 2731
|||||
841 SerSerGlnSerValGlnSerIleArgProProLysArgAlaValSer 857
|||||
2732 AGTTCAAGATCTGTGAGTCTATTCGCTCTCAATATACCGAGACAGTCTC 2781
|||||
857 rLeuAspSerProValSerValGlySerSerProProValLysAsnLys 874
|||||
2782 TCTGGATAGCCCTGTTCTGTGTGCTCAAGTCCCGAGTAAAAAATATNCA 2831
|||||
874 eAlaIlePheProMetLeuProLysGlnProMetLeuGlyGlyAsnProArg 890
|||||
2832 GTGCTTTTCCCATGTATCAAAAGCAACCCATGTGGGAGGAATCAAGA 2881
|||||
891 MetMetAspSerGlnGlnLysArgTyrGlySerSerMetGlyGlyProAsnArg 907
|||||
2882 ATGATGATGATGTCAGAAATTAATGCTCAAGTATG..... 2917
|||||
907 gAsnValThrValThrGlnThrProSerSerGlyAspThrGlyLeuProAla 924
|||||
|||||
2918 ..... GGAGACTGGGGCTTACCAA 2936
|||||
924 snSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyArg 940
|||||
2937 ACTCAAGGCGCGCAGATGGAACCTATGAATTAACCTCATGGGAGAGA 2986
|||||
941 ProGlyGlyAspTyrAsnThrSerLeuProArgProAlaLeuGlyLysG 957
|||||
2987 CCAGGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3036
|||||
957 rIleProThrLeuProLeuArgSerAsnSerIleProGlyValAlaArgPro 974
|||||
3037 TATTCACCATTTGGCTCTTGGCTTAATAGCATTCAGAGTCCGAGACAG 3086
|||||
974 aLleuGlnGlnGlnGlnGlnMetLeuGlnMetArgProGlyGlnIlePro 990
|||||
3087 TATTCACACACACACAGAGATGCTTCAAAAGAGCCCTGGGGAATATCCC 3136
|||||
991 MetGlyMetGlyAlaAsnProTyrGlyGlnAlaAlaAlaSerAsnGlnIle 1007
|||||
3137 ATGGGAATGGGGGCTTAATCCCTATGCGCCAGACAGCATCTTAACCACT 3186
|||||
1007 uGlySerThrProAspGlyMetLeuSerMetGlnGlnValSerHisGlyT 1024
|||||
3187 GGGTTCCGCGCCGATGGCATGTTTCATGGAACAAGTTTCTCATGSCA 3236
|||||
1024 hrGlnAsnArgProLeuLeuArgAsnSerLeuAspAspLeuValGlyPro 1040
|||||
3237 CTCAAAATAGGCTCTTCTTGAAGATCCCTGGAATGATCTTGTGTGGGCA 3286
|||||
1041 ProSerAsnLeuGlnGlyGlnSerAspGlnArgAlaLeuLeuAspGlnIle 1057
|||||
3287 CCTTCCAACTCGAAGCCAGAGTGCAGAAAGACATTAATTTGGACCACT 3336
|||||
1057 uHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuGlnGlnIleAsp 1074
|||||
3337 GCACACTCTTCTCAGCAACACAGATGCCACAGGCTCGAAGAAATTTGCA 3386
|||||
1074 rGAlaLeuGlyIleProGlnLeuValAsnGlnGlnGlnAlaLeuGlnPro 1090
|||||
3387 GAGCTTTGGGCTTCTCGAATCTGTCAATCAAGGAGGCTTGTAGAGGCC 3436
|||||
1091 LysGlnAspAlaPheGlnGlnGlnGlnGlnAlaAlaValMetMetAspGlnIle 1107
|||||
3437 AAACAGGATGCTTTCCAAAGCCCAAGAACAGACATTAATGATGATCAGAA 3486
|||||
1107 saLagLysLeuTyrGlnGlnThrTyrProAlaGlnGlnLysProProMetGln 1124
|||||
3487 GCGAGATTAATATGAGACAGACATACCCAGAGGGGCTCCAAATGCAAG 3536
|||||
1124 LysGlyPheHisLeuGlnGlnGlnGlnSerProSerPheAsnSerMetMetAsn 1140
|||||
3537 GAGGCTTTCACTTCAGGAGCAATCACCATCTTTTAATCTCATATGATGAT 3586
|||||
1141 GlnMetAsnGlnGlnGlnLysAsnPheProLeuGlnGlnGlnMetHisPleArgAl 1157
|||||
3587 CAGATGAACCGCAAGGCAATTTTCTCTCTCCAAAGAAATGACCCACAGGCG 3636
|||||
1157 aaSnIleMetArgProArgThrAsnThrProLysGlnLeuArgMetGlnI 1174
|||||
3637 CAACATATATGAGACCCGAGCAAAACCCCAACCACTTAATGAATGACGCG 3686
|||||
1174 euGlnGlnArgLeuGlnGlnGlnGlnPheLeuAsnGlnSerArgGlnAla 1190
|||||
3687 TTCAGCAGAGGCTCAGAGGCGCAGAGTTTTGAATTCAGAGCGCAGAGCA 3736
|||||
1191 LeuGlnLeuLysMetGlnLysAsnProThrAlaGlyLysAlaAlaValMetArg 1207
|||||
3737 CTTGAATTTGAATGAAGAAACCTTACTGCTGTGTGTGTGTGTGTGTGTGTGT 3786
|||||
1207 gProMetMetGlnProGlnGlnGlnGlnPheLeuAsnAlaGlnMetValAlaG 1224
|||||
3787 GCTATATATGAGCCCAAGAGGCTTTTCTTAATGCTCAAAATGCTGCGCC 3836

```



```

1224 1nArSerArgLluleuSerHisPheArgGlnGlnArgValAla 1240
      |||
3837 AAGCAGAGAGAGAGCTGTAAGTCACTTCGACAAAGAGGGGCT 3886
      |||
1241 MetMetMet.....GlnGlnGlnGlnGlnGlnGlnGln 1253
      |||
3887 ATGATGATGACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCA 3936
      |||
1253 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1270
      |||
3937 GCAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCA 3986
      |||
1270 1aPheSerProProProProValThrAlaSerProSerMetAspLysLeu 1286
      |||
3987 CCTCAGCCACCTCTTAATGAGCTGCTTCCCGCAGATGATGGCTT 4036
      |||
1287 1euAlaGlyProThrMetProGlnAlaProProGlnGlnPheProTyrGln 1303
      |||
4037 TTGGCAGGACCCACAATGCCACAGCTCCTCGCAACAGTTTCCATATCA 4086
      |||
1303 nProAsnTyrGlyMetGlyGlnGlnProAspProAlaPheGlyArgValS 1320
      |||
4087 ACCAATTTATGSAATGGGACCAACACAGATCCAGCTTTGGTCCAGTGT 4136
      |||
1320 eSerProProAsnAlaMetMetSerSerArgMetGlyProSerGlnAsn 1336
      |||
4137 CTAGTCCTCCCAATGATGATGTCTCAAGATGGGTCCCTCCAGAAAT 4186
      |||
1337 ProMetMetGlnHisProGlnAlaAlaSerTleTyrGlnSerSerGlnLe 1353
      |||
4187 CCCATGATGACCAACCCGAGGCTGATCATCTATCATAGTCTCCAGAAAT 4236
      |||
1353 tLysGlyTyrProSerGlyAsnLeuAlaArgAsnSerSerPheSerGlnG 1370
      |||
4237 GAAGGGCTGGCCATCAGGAATTTGGCCGGAACAGCTCTTTCCACAC 4286
      |||
1370 1nGlnPheAlaHisGlnGlyAsnProAlaValTyrSerMetValHisMet 1386
      |||
4287 AGCAGTTTGGCCACCAAGGGAATCTGAGTATAGTATGTGTCACATG 4336
      |||
1387 AsnGlySerSerGlyHisMetGlyGlnMetAsnMetAsnProMetProMe 1403
      |||
4337 AATGGCAGCAGGTGTCATGAGGACAGATGAAATGAAACCCATGGCCAT 4386
      |||
1403 tSerGlyMetProMetGlyProAspGlnLysTyrCys 1415
      |||
4387 GTCGTGCAATGCCCTATGGGTCTGATCAGAAATATCTGC 4423
      |||

seq_name: /cgn2_6/ptoddata/1/pna/US096C_COMB.seq:US-09-652-355-11058

seq_documentation_block:
: Sequence 11058, Application US/09652355
: GENERAL INFORMATION:
: APPLICANT: Shyjan, Andrew W.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
: TITLE OF INVENTION: THEREFOR
: FILE REFERENCE: 1600.1192-001
: CURRENT APPLICATION NUMBER: US/09/652.355
: PRIOR FILING DATE: 2000-08-30
: PRIOR APPLICATION NUMBER: 60/151.136
: PRIOR FILING DATE: 1999-08-30
: NUMBER OF SEQ ID NOS: 11227
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11058
: LENGTH: 7888
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-652-355-11058

```

alignment_scores: 7233.50

length: 1429

```

Ratio: 5.174      Gaps: 3
Percent Similarity: 97.831      Percent Identity: 97.621
Alignment block:
US-09-041-994-2 x US-09-652-355-11058      ..
Align seg 1/1 to: US-09-652-355-11058 from: 1 to: 7888

1 MetSerGlyLeuGlyGluAsnLeuAspProLeuAlaSerAspSerArgly 17
  |||
182 ATGAGTGGATTAGGAAACCTTGGATCCACTGGCAGATGATTCAGAA 231
  |||
17 sArgLysLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGlyG 34
  |||
232 ACGCAAAATGGCATGTGATCTCCAGACAAAGCTCTTACTCTCAATGGG 281
  |||
34 1uLysArgArgArgGluGlnGlnLysSerTyrTleGluLuleuAlaGlu 50
  |||
282 AAAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 331
  |||
51 1eu1leSerAlaAsnLeuSerAsp1leAspAsn1leAsnValLysProAs 67
  |||
332 CTGATATCTGCCAATCTTATGATATGACATTTCAATGTCAAACAGCA 381
  |||
67 pLysCysAla1leLeuLysGluThrValArgGln1leArgGln1leLysG 84
  |||
382 TAAATGTGGCATTTTAAAGAAACAGTAAGACAGATACGTCAAATAAAG 431
  |||
84 1uGlnGlyLysThr1leSerAsnAspAspAspValGlnLysAlaAspVal 100
  |||
432 ACCAAGGAAACCTATTTCCATGATGATGATGATTCCTCAAAACCCATG 481
  |||
101 SerSerThrGlyGlnGlyVal1leAspLysAspSerLeuGlyProLeu 117
  |||
482 TCTTGTACAGGGCAGAGAGTATGATTAAGACTCCTTAGACCCGCTT 531
  |||
117 1uLeuGlnAlaLeuAspGlyPheLeuPheValAlaAsnArgGlnAlaAsn 134
  |||
532 ACTTCAGGATGGATGGATGTTCTTATTTGTGTGATCAGACGGAACA 581
  |||
134 1eValPheValSerGluAsnVal1ThrGlnTyrLeuGlnTyrLysGlnG 150
  |||
582 TTGTATTTTATCAGAAATGTCACACAAATACCTCAATATTAAGCAAG 631
  |||
151 AspLeuValAsnThrSerValTyrAsn1leuHisGluGluAspArgly 167
  |||
632 GACCTGGTTAAACAAAGTGTACAAATATCTTACATGAAGAGACAGAA 681
  |||
167 sAspPheLeuLysAsnLeuProLysSerThrValAsnGlyVal1SerThr 184
  |||
682 GGATTTTCTTAAGAAATTTACAAAAATCTACAGTTAAGAGTTTCTCG 731
  |||
184 hrAsnGluProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200
  |||
732 CAATGAGACCCCAAGACAAACAAACGATATTAATTTGCCGTATGTTG 781
  |||
201 MetLysThrProHisAsp1leLeuGlnAsp1leAsnAlaSerProGlu 217
  |||
782 ATGAACAAACACACATGATATCTGGAAGACATAAAGCCAGAGCTGAAT 831
  |||
217 tArgGlnArgGlyGluThrMetGlnCysPheAlaLeuSerGlnProArg 234
  |||
832 GCGCAGAGATATGAACAAATGACAGTCTTCCCTGCTCCAGCAGG 881
  |||
234 1aMetMetGluGluGlyGluAspLeuGlnSerCysMet1leCysValAla 250
  |||
882 CTATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 931
  |||
251 ArgArgGlyLeuThrGlyGluArgThrPheProSerAsnProGluSerPh 267
  |||
932 CCGCGCATTAACAAGAGAAAGAACATTTCCATCAAAACCTGAGAGCT 981
  |||
267 e1leThrArgHisAspLeuSerGlyLysValAlaValAsn1leAspThrAs 284

```



```

|||||
982 TATTACACACATGATCTTTGAGAAAGTTGTCATATATAGATAAATT 1031
284 eTLeuArgSerSerMetArgProGlyPheGluAspIleIleArgArgCys 300
|||||
1032 CACTGAGATCTTCATGAGAGCGCTGCTTGAAGATATAATCCGAAAGTGT 1081
301 IleGlnArgPhePheSerLeuAsnAspGlyGlnSerTrpSerGlnIleAsr 317
|||||
1082 ATTACAGAGATTTTATAGTCTAAATGATGGCGATGTGTCGCCAGAAAG 1131
317 gHISrYrGlnGlu.....AlaTrpL 324
|||||
1132 TCACATATCAAGAAGTTACCAAGTATGATATTTTCCCAACAGCTTATC 1181
324 euAsnGlyIleHisAlaGluThrProValTyrArgPheSerLeuAlaAspGly 340
|||||
1182 TTTAATGGCCATGCAGAAACCCAGATATCGATTTCTGTTGGCTGATGGA 1231
341 ThrIleValThrAlaGlnThrLysSerLysLeuPheArgAsnProValTh 357
|||||
1232 ACTATATGATCTGCACAGACAAAGACCAATCTTCCGAAATCTGTATAC 1281
357 rAsnAspArgHisGlyPheValSerThrHisPheLeuGlnArgGluGlnA 374
|||||
1282 AATATGATCGACATGGCTTGTCTCAACCCACTTCTCAGAGAGAAACAGA 1331
374 sngLYrTyrArgProAsnProAsnProValGlyGlnGlyIleArgProPro 390
|||||
1332 ATGGATATATGACCAAAACCAAAATCTCTGTGGACAAAGGATTTAGACCACT 1381
391 MetAlaGlyCysAsnSerSerValGlyLysMetSerMetSerProAsnG 407
|||||
1382 ATGGCTGATGCAACAGTTCGGTAGGCGCATGATGTCCGCAAAACCA 1431
407 nGlyLeuGlnMetProSerSerArgAlaTyrGlyLeuAlaAspProSerT 424
|||||
1432 AGGCTTACAGATGCCGAGCAGAGGCGCTATGGCTTGGCAGACCCCTAGA 1481
424 hTrpGlyGlnMetSerGlyAlaArgTyrGlyLysSerSerAsnIleAla 440
|||||
1482 CCACAGGCGACGATGAGTGGAGTAGTATGGGCTTCCAGTAAACATAGCT 1531
441 SerLeuThrProGlyProGlyMetGlnSerProSerSerTyrGlnAsnAs 457
|||||
1532 TCATTGACCCCTGGGCGCAGGCAATGATCAATCATTCTTACCAGAAACA 1581
457 nasnTYrGlyLeuAsnMetSerSerProProHisGlySerProGlyLeuA 474
|||||
1582 CAAGCTATAGGCTCAACATGATGAGTACCCGCCACATGGAGACTCCGTCTTG 1631
474 lApProAsnGlnGlnAsnIleMetIleSerProArgAsnArgGlySerPro 490
|||||
1632 CCCCCAAACCGACGAAATATCATGATTTCTCTCGTATCGTGGAGATGCA 1681
491 LysIleAlaSerHisGlnPheSerProValAlaGlyValHisSerProMe 507
|||||
1682 AAGATAGCCCTCAATCAGTTTCTCTGTTGAGGAGTGCACACTCCCAAT 1731
507 tAlaSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerSerLeuAs 524
|||||
1732 GGCATCTTCTGGCAATACATGGGAGACACAGCTTTTCCACAGACTCTCTCA 1781
524 eTAlaLeuGlnAlaIleSerGlyGlyValGlyThrSerLeuLeuSerThr 540
|||||
1782 GTGCGCTGCAGACCATCATGTAAGGTGTGGGACCTCCCTTTATCTACT 1831
541 LeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIlePh 557
|||||
1832 CTGTCTATCCAGGCCCAAAATTTGATTAATCTCCCAATATGAAATTTAC 1881
557 rGlnProSerLysValSerAsnGlnAspSerLysSerProLeuGlyPheT 574
|||||

```

```

1882 CCAACCAAGTAAAGTAAAGCAATCAGGATTCCAAGAGTCCCTGGGCTTTT 1931
574 YrCYsAspGlnAsnProValAlaGluSerSerMetCysGlnSerAsnArg 590
|||||
1932 ATTCGACCAAAAATCCAGTGGAGAGTTCAATGTGTAGCAAAATATACACA 1981
591 AspHisLeuSerAspLysGluSerLysGluSerSerValGlyValAlaG 607
|||||
1982 GATCACCTCAGTGACAAAGAAAGTAAGAGACAGCTTGAGAGGGGCACA 2031
607 uAsnGlnArgGlyProLeuGluSerLysGlyHisLysLysLeuLeuGlnL 624
|||||
2032 GAATCAAAGGGGTCTTTGGAAGCAAGGTCATATAAAATATCTGCAGT 2081
624 euLeuThrCysSerSerAspAspArgGlyHisSerSerLeuThrAsnSer 640
|||||
2082 TACTTACCTGTTCTTGATGACCGGGGTCATCTCCCTTGACCAACTCC 2131
641 ProLeuAspSerSerCysLysGluSerSerValSerValThrSerProSe 657
|||||
2132 CCCTGATTCAGATGTTAAAGATCTTCTGTATGATGCCAGCCCTTC 2181
657 rGlyValSerSerSerThrSerGlyValAlaSerSerThrSerAsnMetH 674
|||||
2182 TGGAGTCTCTCCTCTACATCTGGAGAGATATCTCTACATCCAAATATGC 2231
674 iSGlySerLeuLeuGlnGlyLysHisArgIleLeuHisLysLeuLeuGln 690
|||||
2232 ATGGGTCACTGTTTACAAAGAAAGCACCGGATTTTGCACAAATGTCGCGAG 2281
691 AsnGlyAsnSerProAlaGlnValAlaLysIleThrAlaGlnAlaThrG 707
|||||
2282 AATGGAAATTCACCACTGATGAGCAAGATTTACTGCAAGAACCACTGG 2331
707 YrLysAspThrSerSerIleThrSerCysGlyAspGlyAsnValValLysG 724
|||||
2332 GAAAGACACACAGACATATAACTTCTGTGGGACGGAATGTTGTCAACG 2381
724 lngLglnLeuSerProLysLysLysGluAsnAsnAlaLeuLeuArgTyr 740
|||||
2382 AGGAGCAGCTAAGTCTTAAGAAAGAGAAATTAATCATCTTGTAGATAC 2431
741 LeuLeuAspArgAspAspProSerAspAlaLeuSerLysGlnLeuGlnPr 757
|||||
2432 CTGCTGACAGGAGATATCTAGTGAATGCATCTTAAGAACTACAGCC 2481
757 oGlnValGlyGlyValAspAsnLysMetSerGlnCysThrSerSerThrI 774
|||||
2482 CCAAGTGGAAAGAGTGGATATAATAATGATGATGCACACACTCCACCA 2531
774 lApProSerSerSerGlnGlyLysAspProLysIleLysThrGluThrSer 790
|||||
2532 TTCTCTGATCTCAAGTCAAGAGAAAGACCTTAAATTTAAGACAGACAACT 2581
791 GluGlnGlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLe 807
|||||
2582 GAAAGAGGATCTGGAGACTTGGATTAATCTTAGATGCTATCTTGGTATGT 2631
807 uThrSerSerAspPheTyrAsnAsnSerIleSerSerAsnGlySerHisL 824
|||||
2632 GACTATCTTGACTTTTCAATATAATTCATATCCTTAATGGTATGATCAGC 2681
824 euGlyThrLysGlnGlnValPheGlnGlyThrAsnSerLeuGlyLeuLys 840
|||||
2682 TGGGAGCTTAAGCAACAGGTGTTTCAAGGAACATAATTTCTGGGTTGAAA 2731
841 SerSerGlnSerValGlnSerIleArgProProLysArgArgAlaValSe 857
|||||
2732 AGTTCAAGTCTGTGAGTCTATGTCTCTCATATTAACCGAGCAAGTGTG 2781
857 rLeuAspSerProValSerValGlySerSerProProValLysAsnIleS 874
|||||
2782 TCTGGATACCTCTGTTTCTGTGGCTCAAGTCTCTCAGTAAAAATATATCA 2831

```

874 eraIaPheProMetLeuProLysGlnProMetLeuGlyGlyAsnProArg 890
 2832 GTGCTTCCCCATGTTACCAAAAGCAACCATGTTGGGTGGAAATCCAGA 2881
 891 MetMetAspSerGlnGluAsnTyrGlySerSerMetGlyGlyProAsnAr 907
 2882 ATGATGGATGTGTCAGAAAAATTATGGCTCAAGTATG..... 2917
 907 GAsnValThrValThrGlnThrProSerSerGlyAspTyrPglYleuProA 924
 2918GGAGACTGGGGCTTACCAG 2936
 924 snserLysAlaGlyArgMetGluProMetLAsnSerAsnSerMetGlyArg 940
 2937 ACTCAAAAGCGCGCAGATGGAACCTATGAATTCMAACTCCATCGGAGAGA 2986
 941 ProGlyLysAspTyrAsnThrSerLeuProArgProAlaLeuGlyGlySe 957
 2987 CCAAGAGAGATTAATTAATCTTCTTACCCAGACCTGCACCTGGGTGGCTC 3036
 957 rIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgProV 974
 3037 TATTCACACATGGCTCTTGGCTTAATAGCATACAGGTGCGAGACAG 3086
 974 aLeuGlnGlnGlnGlnMetLeuGlnMetArgProGlyGlyIlePro 990
 3087 TATTGCAACAGCAGCAGCAGATGCTCAAAATGAGCCCTGGTGAATCCCC 3136
 991 MetGlyMetGlyAlaAsnProTyrGlyGlnAlaAlaAlaSerAsnGlnIle 1007
 3137 ATGGGAATGGGGGCTAATCCCTATGGCCAGACAGCAGCATTTAACCAACT 3186
 1007 uGlySerTyrProAspGlyMetLeuSerMetGluGlnValSerHisGlyT 1024
 3187 GGGTCTCTGGCCGATGSCATGTTGTCCATGAGAACAGTTTCTCATGGCA 3236
 1024 hrgLAsnArgProLeuLeuArgAsnSerLeuAspAspLeuValGlyPro 1040
 3237 CTCAAAATAGGCTCTTCTTGAAGAAATCCCGATGATCTTGTGGGCCCA 3286
 1041 ProSerAsnLeuGlnGlyGlnSerAspGlyuArgAlaLeuLeuAspGlnIle 1057
 3287 CCTTCCAACTCGGAAGCCAGACAGTACGAAAGACATTTATGGACCAACT 3336
 1057 uHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuGlnGluIleAspA 1074
 3337 GCACACTCTTCTCAGCAACACAGATGCCACAGCGCTGGAAGAAATGACA 3386
 1074 rglAlaLeuGlyIleProGluLeuValAsnGlnGlyGlnAlaLeuGluPro 1090
 3387 GAGCTTTGGGCAATTCGAACTTGTCAATCAGGGACAGGCATTAGACCCC 3436
 1091 LysGlnAspAlaPheGlnGlyGlnGlnAlaAlaValMetMetAspGlnIle 1107
 3437 AAACAGATGCTTTCCAAAGCCCAAGAACAGACAGATTAATGATGATCGAA 3486
 1107 saLaGlyLeuTyrGlyGlnThrTyrProAlaGlnGlyProProMetGlnG 1124
 3487 GCGAGGATTTATGACAGACATACCCAGCAGAGGGCTCCATATGCAAG 3536
 1124 LysGlyPheHisLeuGlnGlyGlnSerProSerPheAsnSerMetMetLAsn 1140
 3537 GAGGCTTTCACTCTTCAGGAGCAATCACCATCTTTAACTCATATGATGAT 3586
 1141 GlnMetAsnGlnGlnGlnLysAsnPheProLeuGlnGlyMetHisProArgAl 1157
 3587 CAGTGTGAACGCAAGGCAATTTCTCTCCAAAGGAATGCAACCCAGAGAC 3636
 1157 aaSnIleMetArgProArgThrAsnThrProLysGlnLeuArgMetGlnI 1174
 3637 CAACATCATGAGACCGCGAACAACACCCCAACCACTTAGATGACAGC 3686

1174 euGlnGlnArgLeuGlnGlyGlnGlnPheLeuAsnGlnInSerArgGlnAla 1190
 3687 TTCCAGCAGAGCTGTCAGCGCAGAGCTTTTGAATGAGACCCGACAGCA 3736
 1191 LeuGluLeuLysMetGluAsnProThrAlaGlyGlyAlaAlaValMetArg 1207
 3737 CTTCATTTGAATAATGAAACCCTACTGCTGCTGCTGCTGCTGCTGCTGCT 3786
 1207 gProMetMetGlnProGlnGlnGlnPheLeuAsnAlaGlnMetValAlaG 1224
 3787 GCCTATGATGACAGCCCGACAGAGGCTTTTCTTAATGCTCAATGGTCCCC 3836
 1224 LnaArgSerArgGluLeuLeuSerHisHisPheArgGlnGlnArgValAla 1240
 3837 AACCCAGCAGAGAGCTCTTAATGTCATCTCCGACACAGAGAGGTGGCT 3886
 1241 MetMetMet.....GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1253
 3887 ATGATGATGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 3936
 1253 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1270
 3937 GCACACACAGACACAGACACAGACACAGACACAGACACAGACACAGCAG 3986
 1270 LapeSerProProProAsnValThrAlaSerProSerMetAspGlyLeu 1286
 3987 CTTTCACCCCACTCTTAATGTGACTCTTCCCGCAGCATGGATGGGCTT 4036
 1287 LeuAlaGlyProThrMetProGlnAlaProProGlnGlnPheProTyrGly 1303
 4037 TTGGCAGAGCCCAACATGCCACAGCTCTTCCGCAACAGTTTCCATATCA 4086
 1303 nProAsnTyrGlyMetGlyGlnGlnProAspProAlaPheGlyArgValS 1320
 4087 ACCAAATTTATGAAATGGGACAAACACAGATCCAGCTTGGTGGAGTGT 4136
 1320 eSerProProAsnAlaMetMetSerSerArgMetGlyProSerGlnAsn 1336
 4137 CTAGTCTCTCCAAATGCAATGATGTGTCAGAAATGGGTCTCCCTCCACAAT 4186
 1337 ProMetMetGlnHisProGlnAlaAlaSerIleTyrGlnSerSerGlnIle 1353
 4187 CCCATGATGACACACCGCAGCTGCATTCATCATCATAGTCTAGAAAT 4236
 1353 tLysGlyTyrProSerGlyAsnLeuAlaArgAsnSerSerPheSerGln 1370
 4237 GAAGGCTGGCCATCAGAAATTTGGCCAGGACAGCTCTTTCCAGC 4286
 1370 lngInPheAlaHisGlnGlyAsnProAlaValTyrSerMetValHisMet 1386
 4287 AGCAGTTTGGCCACAGGGGAATCTTCAGTGTATGATGTGCACATG 4336
 1387 AsnGlySerSerGlyHisMetGlyGlnMetLAsnMetAsnProMetPPro 1403
 4337 AATGCGACAGATGCTCACATGGACATGACATGAACCCCATGGCCAT 4386
 1403 tSerGlyMetProMetGlyProAspGlnLysTyrCys 1415
 4387 GTCTGCGATGCTTATGGGTCTTCATCAGAAATACTGC 4423
 seq_name: /cgn2_6/plodata/1/pna/us096D_COMB.seq:us-09-663-693-1028
 seq_documentation_block:
 ; Sequence 1028, Application us/09663693
 ; GENERAL INFORMATION:
 ; APPLICANT: Geating, David P.
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 ; FILE REFERENCE: 1600.1195-001
 ; CURRENT APPLICATION NUMBER: US/09/663,693
 ; CURRENT FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: 60/154,986
 ; PRIOR FILING DATE: 1999-09-21

NUMBER OF SEQ ID NOS: 1340
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 1028
LENGTH: 7888
TYPE: DNA
ORGANISM: Homo sapiens
US-09-663-693-1028

alignment_scores:
Quality: 7233.50 Length: 1429
Ratio: 5.174 Gaps: 3
Percent Similarity: 97.831 Percent Identity: 97.621

alignment_block:
US-09-041-994-2 x US-09-663-693-1028

Align seg 1/1 to: US-09-663-693-1028 from: 1 to: 7888

```
1 MetSerGlyLeuGlyLysAsnLeuAspProLeuAlaSerAspSerArgly 17
|||||
182 ATGAGTGGATTAGGAGAAACCTGGATCCAGTGGCCAGTATTCACGAAA 231
17 sarGlyLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGly 34
|||||
232 ACCCAATTGCCATGTGATCTCAGACACAGCTTACCTGCGAGTGTG 281
34 lulysArgArgArgGluGlnGlnSerLysTyrIleGluGlnLeuAlaGlu 50
282 AAAACGCGAGCGGACGAGAAAGTAAATATATGAAAGATTGGCTGAG 331
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67
|||||
332 CTGATATCTCCCAATCTTACTGATATGACAAATTCATGTCAAACGAG 381
67 pLysCysAlaIleLeuLysGlnThrValArgGlnIleArgGlnIleLysG 84
|||||
382 TAATATGGCATTTTAAAGAAACAGTAACACAGATACGTCMAATTAAG 431
84 LulGlnGlyLysThrIleSerAsnAspAspValGlnLysAlaAspVal 100
432 AGCAAGGAAACAACTATTCATGATGATGATGATCAAAAAGCCGATGTA 481
101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeu 117
|||||
482 TCTTCTACAGGCGAGGAGTTATGATTAAGACTCCCTTAGACCGCTTTP 531
117 ulGlnAlaLeuAspGlyPheLeuPheValIleAsnArgGlnAlaAsnI 134
532 ACTTCAGGCATTGGATGCTTCCATTTGTGCTGAATCGAGACGAGAA 581
134 leValPheValSerGlnAsnValThrGlnTyrLeuGlnTyrLysGln 150
582 TTTGATTTGATGAGAAATGTCACACAAATACCTGCAATATTAAGCAAG 631
151 AspleuValAsnThrSerValTyrAsnIleLeuHisGluGlnAspArgly 167
632 GACCTGGTTAACACAAAGTGTTCACAAATATCTTACATGAAGAGCAAAA 681
167 saSPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrp 184
682 GGATTTTCTTAAGATTTTACCAAAATCTACAGTTAATGAGATTCCCGA 731
184 hrAsnGlnProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200
732 CAATATGAGCCCAAGCAAAAGCAAAAGCATATTATTCGCCGATGTGT 781
201 MetLysThrProHisAspIleLeuGlnAspIleAsnAlaSerProGlu 217
782 ATGAAAACACACATGATATTCGGAAGACATTAACGCCAGTCTGTAAT 831
217 LArgGlnArgGlyLysThrMetGlnCysPheAlaLeuSerGlnProArg 234
|||||
```

```
832 GCGCCAGATATGAAACAAATGCAAGTCTTGGCCCTGCTCAGCCAG 881
234 lamMetGluGlnGlyLysAsnLeuGlnSerCysMetIleCysValAla 250
882 CTATGATGAGGAGGAGGAGATTTGCAATCTTGTATGATCTGTGTGCA 931
251 ArgArgIleThrThrGlyLysArgThrPheProSerAsnProGlySer 267
932 CGCCGATTTACTACGAGAAAGCAATTTCCATCAAAACCTTGAGACTT 981
267 eileThrArgHisAspleuSerGlyLysValIleAsnIleAspThrAsn 284
982 TATTACGACATCATCTTTCAGGAAAGTGTCTCAATATGATGATCAAT 1031
284 erLeuArgSerSerMetArgProGlyPheGlnAspIleIleArgArgCys 300
1032 CACTGAGATCTCCATGAGCGCTTGCTTGAAGATATATATCCGAAGTGT 1081
301 IleGlnArgPhePheSerLeuAsnAspGlyGlnSerThrSerGlnLys 317
1082 ATTCAGAGATTTTATGTTAAATGATGGCAGTCATGATGTCGAGAAAC 1131
317 gHisTyrGlnGlu.....AlaTyrL 324
1132 TCACATACAAAGAACTTACCACTGATGGGATATTTTCCCAACAGCTTAT 1181
324 euAsnGlnHisAlaGlnThrProValTyrArgPheSerLeuAlaAspGly 340
1182 TTATATGCCATGCGAGAAACCCAGTATATGATCTGCTGGCGATGAGA 1231
341 ThrIleValThrAlaGlnThrLysSerLysLeuPheArgAsnProVal 357
1232 ACTATAGTACTGCACAGACAAAAAGCAAACTCTCGAAATCCTGTGAT 1281
357 rAsnAspArgHisGlyPheValSerThrHisPheLeuGlnArgGlnIle 374
1282 AAATGATCGACATGCTTTGTCTCAACCCACTTCTTCAGAGAAACAGA 1331
374 snGlyTyrArgProAsnProAsnProValGlnGlnLysIleArgPro 390
1332 ATGATATATGACCAAAACCAATCTGTTGACAAAGGATTAAGCACACT 1381
391 MetAlaGlyCysAsnSerSerValGlyLysMetSerMetSerProAsn 407
1382 ATGGCTGATGCAACACTGCTGATGCGGCAATGATGTCGCCAAACCA 1431
407 nGlyLeuGlnMetProSerSerArgAlaTyrGlyLeuAlaAspProSer 424
1432 AGGCTTACAGATGCGGAGACAGAGGCTATGCTTGGACGCCCTAGCA 1481
424 hrThrGlyGlnMetSerGlyAlaArgTyrGlyLysSerSerAsnIleAla 440
1482 CCACAGGCGAGATGAGTGGAGCTGATGAGGCTTCACATACATAGCT 1531
441 SerLeuThrProGlyProGlyMetGlnSerProSerSerLysGlnAsn 457
1532 TCATTGACCCCTGGCCAGGATGCAATCACCATCTTCTTACCAAGAA 1581
457 nasTyrGlyLeuAsnMetSerSerProProHisGlySerProGlyLeu 474
1582 CAATATAGGCTCAACATGATGATGAGCCCAACATGAGGAGTCTGTG 1631
474 laProAsnGlnGlnAsnIleMetIleSerProArgAsnArgLysPro 490
1632 CCCCAAAACGACAGAAATATCATGATTTCTCTCTAATCTGTGGAGTCA 1681
491 LysIleAlaSerHisGlnPheSerProValAlaGlyValHisSerPhe 507
1682 AAGTATGCTTCACTCAAGTTTCTCTGTTGACAGTGTGACATCTCCAT 1731
507 talAserSerGlyAsnThrGlyAsnHisSerPheSerSerSerSerLeu 524
1732 GGCATCTTTCGCAATCTGGGAACACACACTTTTCCAGCAGCTCTCTCA 1781
```

524 eRLAlaLeuGlnAlaIleSerGluGlyValGlyThrSerLeuSerThr 540
 |||||
 1782 TGCCCTGCAACCCATGAGTGGGGGACTCCCTTTATCTACT 1831
 541 LeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleTh 557
 |||||
 1832 CTGTATTCACAGGCCCAATTGGATTACTCTCCAAATATGATATTTC 1881
 557 rGlnProSerLysValSerAsnGlnAspSerLysSerProLeuGlyPheT 574
 |||||
 1882 CCAACCAAGTAAGTAAGCAATCAGATTCCAAGAGTCTCTGGGCTTTT 1931
 574 yrcYAspGlnAsnProValGluSerSerMetCysGlnSerAsnSerArg 590
 |||||
 1932 ATTGGACCAAAATCCAGTGGAGAGTTCATGTCAGTCAATATACAGA 1981
 591 AspHisLeuSerAspLysGluSerLysGluSerSerValGluGlyIleG 607
 |||||
 1982 GATCACCTCAGTGACAAAGAAAGTAAGAGACAGCTTGAAGGGGCA 2031
 607 uAsnGlnArgGlyProLeuGluSerLysGlyHisLysLysLeuLeuGln 624
 |||||
 2032 GAATCAAAAGGGGCTTGTGAAAGCAAGGTCTATAAAATTAATCTCAG 2081
 624 euLeuThrCysSerSerAspAspArgGlyHisSerSerLeuThrAsnSer 640
 |||||
 2082 TACTTACCTGTTCTTCTGATGACCGGGGTATTCCTCTTGACCAATCC 2131
 641 ProLeuAspSerSerCysLysGluSerSerValSerValThrSerProse 657
 |||||
 2132 CCCCTGATTCAAAGTTGTAAGAAATCTTGTAGTGCACCAAGCCCTTC 2181
 657 rGlyValSerSerSerThrSerGlyValSerSerThrSerAsnMetH 674
 |||||
 2182 TGGAGTCTCTCTCTACATCTGAGAGAGATATCCCTACATCCAAATATGC 2231
 674 lSGlySerLeuLeuGlnGlyHisArgIleLeuHisLysLeuLeuGln 690
 |||||
 2232 ATGGGCTACTGTTACAAAGAAAGCACCGGATTTGCACAAGTTGCGCAG 2281
 691 AsnGlyAsnSerProAlaGluValAlaLysIleThrAlaGlnAlaThrG 707
 |||||
 2282 AATGGGAATTCACACAGCTAGGTAAGCAAGATTACTGCAGAAAGCCACTGG 2331
 707 yLysAspThrSerSerIleThrSerCysGlyAspGlyAsnValIleLysG 724
 |||||
 2332 GAAAGACACACAGATATAACTTCTTGGGGACGGAATGTTCACAACG 2381
 724 lngGlnLeuSerProLysLysGlyAsnAsnAlaLeuLeuArgIle 740
 |||||
 2382 AGGAGCAGCTAAGTCTTAAGAAAGAGAAATATGCACCTCTTGAATAC 2431
 741 LeuLeuAspArgAspAspProSerAspAlaLeuSerLysGluLeuGlnPr 757
 |||||
 2432 CTGCTGGACAGGAGATGATCTAGTCATGCACTCTCTAAAGAACTACAGCC 2481
 757 oGlnValGluGlyValAspAsnLysMetSerGlnCysThrSerSerThrI 774
 |||||
 2482 CCAAGTGGAAAGAGTGAATTAATTAATGATCAGTGCACAGCTCCACCA 2531
 774 leProSerSerSerGlnGlyLysAspProLysIleLysThrGluThrSer 790
 |||||
 2532 TTCTTAAGCTCAAGTCAAGAAAGACCCCTTAATTAAGACAGACAACT 2581
 791 GluGlnGlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLe 807
 |||||
 2582 GAAGAGGAGATCTGGAGACTTGATATCTAGATGCTATTCTTGGTGATCT 2631
 807 uThrSerSerAspPheTyrrAsnAsnSerIleSerSerAsnGlySerHisL 824
 |||||
 2632 GACTAGTCTGACTTTTACAAATTAATTCATATCCCTCAATGCTAGTCATC 2681

824 euGlyThrLysGlnGlnValPheGlnGlyThrAsnSerLeuGlyLeuLys 840
 |||||
 2682 TGGGGACTAAGCAACAGGTGTTCAAGGAACTAATCTCTGGGTTGAAA 2731
 841 SerSerGlnSerValGlnSerIleArgProProTyrrAsnArgAlaValSe 857
 |||||
 2732 AGTTCCAGCTCTGTGCACTCTATTCGTCCTCCATATTAACCGAGCACTGTC 2781
 857 rLeuAspSerProValSerValGlySerSerProProValLysAsnIleS 874
 |||||
 2782 TCTGGATACCCCTGTTCTGTGGCTCAAGTCTCCAGTAATAAAATATCA 2831
 874 eRLaPheProMetLeuProLysGlnProMetLeuGlyLysAsnProArg 890
 |||||
 2832 GTGCTTTCCTCATGTTTACCAAGCAACCCATGTGTGGTGGAAATCCAGA 2881
 891 MetMetAspSerGlnGluAsnTyrrGlySerSerMetGlyGlyProAsnAr 907
 |||||
 2882 ATGATGGATAGTCAAGAAATTAATGCTCAAGTATG. 2917
 907 gAsnValThrValThrGlnThrProSerSerGlyAspTrpGlyLeuPro 924
 |||||
 2918GGAGACTGGGCTTACCAA 2936
 924 snSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyArg 940
 |||||
 2937 ACTCAAAAGCCGGCAGAAATGGAACCTATGAATTCAAATCTCAGGAGACA 2986
 941 ProGlyGlyAspTyrrAsnThrSerLeuProArgProAlaLeuGlyLysE 957
 |||||
 2987 CCAAGAGGAGATTAATTAATCTTATACCACAGCTGCACCTGGTGCTC 3036
 957 rIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgProV 974
 |||||
 3037 TATTCACATGTCCTCTCGGTGTAATAGCATACAGAGTGGAGACAGCAG 3086
 974 allLeuGlnGlnGlnGlnMetLeuGlnMetArgProGlyLulIlePro 990
 |||||
 3087 TATTTGCACAGCAGCAGCAGATGCTTCATATGAGGCTGTGTAATCCCC 3136
 991 MetGlyMetGlyAlaAsnProTyrrGlyGlnAlaAlaSerAsnGlnLe 1007
 |||||
 3137 ATGGGAATGGGGCTAATCCTATGCGCAAGCAGCATCTTAACCAACT 3186
 1007 uGlySerTrpProAspGlyMetLeuSerMetGluGlnValSerHisGlyT 1024
 |||||
 3187 GGGTTCCTGGCCCGATGGCATGTTGCCATGGAACAAGTTCTCATGGCA 3236
 1024 hrgLlnAsnArgProLeuLeuArgAsnSerLeuAspLeuValGlyPro 1040
 |||||
 3237 CTCAAATATGGCCTCTTCTTAAGAAATCCCTGATGATCTTGTGGCCA 3286
 1041 ProSerAsnLeuGluGlyGlnSerAspGluArgAlaLeuLeuAspGlnLe 1057
 |||||
 3287 CCTTCCAACTGGAAAGCCAGAGTCAAGAAAGCAATTAATGACCAAGCT 3336
 1057 uHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuGlnGluIleAspA 1074
 |||||
 3337 GCACACTTCTCAGAACACAGATGCCACAGGCGCTGGAAAGAAATTGACA 3386
 1074 rGAlaLeuGlyIleProGluLeuValAsnGlnGlyGlnAlaLeuGluPro 1090
 |||||
 3387 GAGCTTTGGGCATTCCTGAACCTTGCAATCAAGGAGCAGCATTAAGGCC 3436
 1091 LysGlnAspAlaPheGlnGlyGlnGlnAlaValMetMetAspGlyL 1107
 |||||
 3437 AAACAGGATCTTTCAGAGGCCAAGAGACAGCAGTAATGATGATCAGAGA 3486
 1107 sAlaGlyLeuTyrrGlyGlnThrTyrrProAlaGlnGlyProPromeGln 1124
 |||||
 3487 GGCAGGATTAATGACAGACATACCAAGCAGCAGGAGGCTCCCAATGCAAG 3536
 1124 lYglyPheHisLeuGlnGlnGlnSerProSerPheAsnSerMetMetAsn 1140

```

3537 GAGCTTTCATCTTCAGGACAAATCACCATCTTTAACTCATATCAAT 3586
1141 GlnMetAsnGlnGlnGlnPheProLeuGlnGlnMetHisProArgAl 1157
3587 CAGATGAACAGCAAGCAATTTTCTCTCCAGAAATGACCCAGCAGC 3636
1157 aAsnIleMetArgProArgThrAsnThrProGlnLeuArgMetGln 1174
3637 CAACATCATGAGACCCCGGACAAACACCCCAACCACTTAAGATGACG 3686
1174 eugGlnArgLeuGlnGlnGlnPheLeuAsnGlnSerArgGlnAla 1190
3687 TTCAGCAGAGGCTCCAGGCGCAGAGTTTGAATCAGAGCGCAGAGCA 3736
1191 LeuGlnLeuLysMetGluAsnProThrAlaGlyAlaAlaValMetAr 1207
3737 CTTGAATTTGAAATGAAACCCCTACTGCTGTGCTGCTGCGTGAATGAG 3786
1207 gProMetMetGlnProGlnGlnGlnPheLeuAsnAlaGlnMetValAla 1224
3787 GCCATGATGAGCCGCCAGCGGGGTTTCTTAATGCTCAATVGGTCCGCC 3836
1224 InatGSerArgGlnLeuLeuSerHisHisPheArgGlnArgValAla 1240
3837 AACGCACAGAGAGACTGCTAAGTCATCATCTCCACAAACAGAGGTGCT 3886
1241 MetMetMet.....GlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1253
3887 ATGATGATGATGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 3936
1253 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1270
3937 GCACAAACAGCAACAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAG 3986
1270 IapHeserProProAsnValThrAlaSerProSerMetAspGlyLeu 1286
3987 CTTTCAGCCCACTCTTAATGTGACTGTTCCCCAGCATGATGGGCTT 4036
1287 LeuAlaGlyProThrMetProGlnAlaProProGlnGlnPheProTyrG 1303
4037 TTGGCAGAGACCACAAATGCCCAAGCTCTCCGCAACAGTTTCCATATCA 4086
1303 nProAsnTyrGlyMetGlyGlnGlnProAspProAlaPheGlyArgVal 1320
4087 ACCAAATTTAGAAATGGACAAACAGATCCAGCCTTGTGTGAGTGT 4136
1320 eRserProProAsnAlaMetMetSerSerArgMetGlyProSerGlnAsn 1336
4137 CTACTCCGCCCAATGCAATGATGTGCTCAAGATGGGTCCCTCCAGAT 4186
1337 ProMetMetGlnHisProGlnAlaAlaSerIleTyrGlnSerSerGlu 1353
4187 CCCATGTATGCAACACCCGACAGCTGCATCATCTATCATAGTCTCAGAA 4236
1353 tLysGlyTyrProSerGlyAsnLeuAlaArgAsnSerSerPheSerGln 1370
4237 GAAGGCTGGCCATCAGGAAATTTGGCCAGGACAGCTCTTTTCCAGC 4286
1370 InGlnPheAlaHisGlnGlnGlnProAlaValTyrSerMetValHisMet 1386
4287 AGCAGTTTGGCCACACAGGGAATCCTGCAGTATAGTATGTGTGACACTG 4336
1387 AsnGlySerSerGlyHisMetGlyGlnMetLysMetAsnProMetProme 1403
4337 AATGGCAGCACTGTGTACATGTGAGCAGATGAACATGAACCCCATGCCAT 4386
1403 tSerGlyMetProMetGlyProAspGlnLysTyrCys 1415
4387 GTCTGGCATGCTTATGGGTCTGTGATCAGAAATTAAGTGC 4423
seq_name: /cgn2_6/ptodata/1/pna/US096E.COMB.seq:US-09-698-014-5135

```

```

seq documentation block:
; Sequence 5135, Application US/09698014
; GENERAL INFORMATION:
; APPLICANT: Geating, David P.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Kingsbury, Gillian A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2014-001
; CURRENT APPLICATION NUMBER: US/09/698,014
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 60/162,363
; NUMBER OF SEQ ID NOS: 6098
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 5135
; LENGTH: 7888
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-698-014-5135

alignment_scores:
Quality: 7233.50 Length: 1429
Ratio: 5.174 Gaps: 3
Percent Similarity: 97.831 Percent Identity: 97.621

alignment_block:
US-09-041-994-2 x US-09-698-014-5135 ..
Align seg 1/1 to: US-09-698-014-5135 from: 1 to: 7888

1 MetSerGlyLeuGlyGluAsnLeuAspProLeuAlaSerAspSerArgly 17
182 ATGATGATGATGAGGAAACCTTGATCCAGTCCAGTCAATGATCAAGAAA 231
17 sArgLysLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGly 34
232 AGCAAAATGCCATGTGATCTCCAGCAAGGCTTAACTGCTGAGTG 281
34 IuLysArgArgArgGluGlnGlnLeuSerLysTyrIleGluGluLeuAlaGlu 50
282 AAAAAGCAGAGCGGAGAGCAAGTAATATATGAAGATTTGGCTGAG 331
51 LeuIleSerAlaAsnLeuSerAspIleAspAsnPheAsnValLysProAs 67
332 CTGATATCTGCCAATCTTAGATATGACAAATTTCAATGTCAATGCCAACA 381
67 pLysCysAlaIleLeuLysGluThrValArgGlnIleArgGlnIleLysG 84
382 TAAATGTGCGATTTTAAAGGAAACGTAAGACAGATGCTAAATTAANG 431
84 IuGlnGlyLysThrIleSerAsnAspAspValGlnLysAlaAspVal 100
432 AGCAAGGAAAAAATATTTCCAAATGATGATGATGATCAAAAAGCCGATGTA 481
101 SerSerThrGlyGlnGlyValIleAspLysAspSerLeuGlyProLeu 117
482 TCTTCTACAGGCGAGGAGTATTGATTAAGATCCCTTAGACCGCTTTT 531
117 uLeuGlnAlaLeuAspGlyPheLeuPheValAlaAsnArgGlnAlaAsn 134
532 ACTTCAGCATTTGGATGTTTCCATTTTGTGTGATGATGATGATGATGATG 581
134 IeValPheValSerGluAsnValThrGlnTyrLeuGlnTyrLysGlnGlu 150
582 TTGATATTGTATCAAGAAATCTCACACAAATACGCAATATATTAACAGAG 631
151 AspLeuValAsnThrSerValTyrAsnIleLeuHisGluGluAspArgly 167
632 GACCTGTGTTAACACAAAGTGTATCAATATCTTACATGTAAGAAAGCAAAA 681
167 sAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTrp 184

```

```

|||||
682 GGATTTCTTAGAATTTACAAAATCTACAGTTAATGAGTTTCTCGA 731
184 hrAsnGluProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200
732 CAAATGAGACCCCAAGACAAAAGCCATATCTTAAATGCGGTATGTTG 781
201 MetLysThrProHisAspIleLeuGlnAspIleAsnAlaSerProGluLe 217
782 ATGAAAAACCAATGATATCTTGGAAGACATAAAGCCAGTCCTGAAAT 831
217 TArgGlnArgGluThrMetGlnCysPheAlaLeuSerGlnProArg 234
832 GGGCCAGAGATATGAACAATGCAGTGTTCCTTCCTCAGCCACGAG 881
234 lAmMetGlnGlnGlyLysAspLeuGlnSerCysMetIleCysValAla 250
882 CTATGATGAGAGAGGGAAGATTGCAATCTTGATGATCTGTGGGA 931
251 ArgArgIleThrThrGlyLysArgThrPheProSerAsnProGluSerPh 267
932 CCGCGCATTAACAAGAGAAAACATTTCCATCAAAACCTGAGAGCTT 981
267 eIleThrArgHisAspLeuSerGlyLysValAlaAsnIleAspThrAsn 284
982 TATTACCAACATGATCTTTCAGGAAGGTTGTCAATATGATACAAAT 1031
284 eLeuArgSerSerMetArgProGlyPheGlnAspIleIleArgCys 300
1032 CACTGAGATCTTCATGAGGCGCTTGGAAGATTAATCCGAAGCTGT 1081
301 lIleGlnArgPhePheSerLeuAsnAspGlyGlnSerTrpSerGlnLys 317
1082 ATTCAAGAGATTTTAAATGATGAGGTCATGGTCCCAAGAAAG 1131
317 gHisTyrgInGlu.....AlaTyrl 324
1132 TCACATACAAAGAGTTACCAAGTATGATTTTCCCAACAGCTTATC 1181
324 euAsnGlyHisAlaGluThrProValTyrrArgPheSerLeuAlaAspGly 340
1182 TTAATAGGCATGCAGAAACCCAGATATCGATTCTGTTGGCTGATGGA 1231
341 ThrIleValThrAlaGlnThrLysSerLysLeuPheArgAsnProValh 357
1232 ACTATAGTACTGCACAGACAAAACAACTTCGGAATCCGTATAC 1281
357 rAsnAspArgHisGlyPheValSerThrHisPheLeuGlnArgGluGln 374
1282 AATATGATGACATGGCTTGTCTCAACCCACTTCCTTCAGAGAAACAGA 1331
374 snGlyTyrrArgProAsnProAsnProValGlyGlnGlyIleArgProPro 390
1332 ATGGATATAGACCAAAACCAATTCCTGTGGACAAGGATTAGACACCT 1381
391 MetAlaGlyCysAsnSerSerValGlyGlyMetSerMetSerProAsnAl 407
1382 ATGGCTGGATGCAACAGTTCGGTAGGCGCATGAGATATCGCCAAACCA 1431
407 nGlyLeuGlnMetProSerSerArgAlaTyrrGlyLeuAlaAspProSer 424
1432 AGGCTTACAGATGCCGAGCAGAGGCCCTATGGCTTGGCAGACCCCTAGA 1481
424 hrThrGlyLysIleMetSerGlyAlaArgTyrrGlyLysSerAsnIleAla 440
1482 CCACAGGGCAGATGAGTGAAGTATGAGGGGTTCCAGTAACATAGCT 1531
441 SerLeuThrProGlyProGlyMetGlnSerProSerSerTyrrGlnAsnAs 457
1532 TCATTGACCCCTGGGCGCAGCATGCAATACCATCTTCTTACCAAAACA 1581
457 nAsnTyrrGlyLeuAsnMetSerSerProHisGlySerProGlyLeuAla 474
|||||

```

```

1582 CAACCTTAGCGTCAACATGATAGCCCCCACATGGGAGTCTCTGCTTC 1631
474 lAProAsnGlnGlnAsnIleMetIleSerProArgAsnArgLysSerPro 490
1632 CCCAAACACGACAGATATCATGATTTCTCTCGTAACTGTGGAGTCCA 1681
491 LysIleAlaSerHisGlnPheSerProValAlaGlyValHisSerProHe 507
1682 AAGATAGCCTCAATCAGTTTCTCTGTTCGACAGGTCTCCTCCCAT 1731
507 lAlaSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerLeuAs 524
1732 GGCATTTCTGGCAATACCTGGGAACACAGCTTTTCCACACTCTCTCA 1781
524 eAlaLeuGlnAlaIleSerGlnGlyValGlyThrSerLeuLeuSerThr 540
1782 GTGCCCTGCAAGCATCAGTGAAGGTGTGGGACTTCCTTTATCTACT 1831
541 LeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleTh 557
1832 CTGTCAATCACAGGCCCAAAATTGATTACTCTCCCAATATGAATATAC 1881
557 rGlnProSerLysValSerAsnGlnAspSerLysSerProLeuGlyPhe 574
1882 CCACCAAGTAAAGTAAAGCAATCAGGATTCCAAGACTCTCTGGGCTTT 1931
574 yrcysAspGlnAsnProValGlnSerSerMetCysGlnSerAsnSerArg 590
1932 ATTCGACCAAAATCCAGTGAAGATTCAATGTCAGTCAATATGACAGA 1981
591 AspHisLeuSerAspLysGlnSerLysGlnSerSerValGlnGlyAlaG 607
1982 GATCACCTCAGTACAAAGAAAGTAAAGAGACAGAGTGGAGGGGACAA 2031
607 uAsnGlnArgGlyProLeuGlnSerLysGlyHisLysLysLeuLeuGln 624
2032 GAATCAAAAGGCTCTTTGGAAGCAAAAGCTATATAAAATTAATACGCA 2081
2082 TACTTACCTGTTCTTCTGATGACCGGGGTCAATTCCTTGACCAACTCC 2131
624 euLeuThrCysSerSerAspAspArgGlyHisSerSerLeuThrAsnSer 640
641 ProLeuAspSerSerCysLysGlnSerSerValSerValThrSerProSe 657
2132 CCCCTAGATTCAAGTTGTAAGAAATCTTGTAGTGTACCAAGCCCTTC 2181
657 rGlyValSerSerSerThrSerGlyValAlaSerSerThrSerAsnMet 674
2182 TGGAGTCTCTCCTCTACATCTGGAGAGTATCTCTACATCCAAATATGC 2231
674 lAsGlySerLeuLeuGlnGlnLysHisArgIleLeuHisLysLeuLeuGln 690
2232 ATGGGTCACTGTTACAAAGAAAGACCCGATTTTGCACAAATGGTGCAG 2281
691 AsnGlyAsnSerProAlaGluValAlaLysIleThrAlaGlnAlaThrGl 707
2282 AATGGGAATTCACCAAGCTGAGTGAAGCAAGATTACTGCAGAAAGCACTGG 2331
707 yLysAspThrSerSerIleThrSerCysGlyAspGlyAsnValValLysG 724
2332 GAAAGACACACAGATATTAATCTTTGTGGGACACGAAATGTTGTCAAGC 2381
724 lngLglnLeuSerProLysLysGlyAsnAsnAlaLeuLeuArgTyrr 740
2382 AGGAGCAGCTTAAGTCTTAAGAAAGAAAGAAATTAATGCAATCTTGTAGATC 2431
741 LeuLeuAspArgAspAspProSerAspAlaLeuSerLysGlnLeuGlnPr 757
2432 CTGCTGACAGGAGATGATCTAGTGTGACACTCTTAAGAAACTTCAGAGC 2481
757 oGlnValGlnGlyValAspAsnLysMetSerGlnCysThrSerSerThrI 774
2482 CCAAGTGAAGAGTGCATTAATAATGATGATGTCAGTCCACCACTCCACA 2531

```

774 leProSerSerSerGlnGlnIuLysAspProLysIleLysThrGluThrSer 790
 |||||
 2532 TTCTACGTCAGAGTCAAGAGAAAGACCTTAATAATTAAGACAGACCAAGT 2581
 791 GlnGlnGlnSerGlyAspLeuAspAsnLeuAspAlaIleGlnGlyAspLe 807
 |||||
 2582 GAAGAGGATGTGAGACTTGATATATCTAGATCTATCTTGCGTGAATCT 2631
 807 uThrSerSerAspPheTyrrAsnAsnSerIleSerSerAsnGlySerHisL 824
 |||||
 2632 GACTAGTCTCACTTTACAAATTAATCCATATCCCTCAAAATGATGATC 2681
 824 euGlyThrGlnGlnIleValPheGlnGlnIleThrAsnSerLeuGlyLeu 840
 |||||
 2682 TGGGAGCTAACCAAGAGTGTTCAGAGAACTAATCTCTGGGTTTGAA 2731
 841 SerSerGlnSerValGlnSerIleArgProProTyrrAsnArgAlaValSe 857
 |||||
 2732 AGTTCACAGTCTGTGAGTGTATCTGCTCTCCATATAACGACGAGTCTC 2781
 857 rLeuAspSerProValSerValGlySerSerProProValLysAsnIleS 874
 |||||
 2782 TCTGGAAGACCCCTGTTCTGTCGCTCAAGTCTCCAGTAAATAATATCA 2831
 874 eAlaPheProMetLeuProLysGlnProMetLeuGlyLysAsnProArg 890
 |||||
 2832 GTGCTTTCCTCCATGTATCCAAAGCAACCCATGTGGGGGAATTCAGAA 2881
 891 MetMetAspSerGlnGlnAsnTyrrGlySerSerMetGlyGlyProAsnAr 907
 |||||
 2882 ATGATGGATGATGACGAAATAATATGGCTCAAGTATG..... 2917
 907 gAsnValThrValThrGlnThrProSerSerGlyAspTrpGlyLeuProA 924
 |||||
 2918GGAGACTGGGGCTTACCAA 2936
 924 snSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyArg 940
 |||||
 2937 ACTCAAAAGCGCGCAGATGAGAACCTTAAGAAATTAACCTCATGGGAGA 2986
 941 ProGlyLysAspTyrrAsnThrSerLeuProArgProAlaLeuGlyLysE 957
 |||||
 2987 CCAGAGAGAGATATATATCTTCTTACCCAGACTGCTGAGGCTGCTC 3036
 957 rIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgProV 974
 |||||
 3037 TATTCACCAATGGCTCTTCGCTTAATAGCATACAGAGTGCAGAGCCAG 3086
 974 alLeuGlnGlnGlnGlnMetLeuGlnMetLeuArgProGlyGluIlePro 990
 |||||
 3087 TATTCACCAACAGCAGCAGATGCTTCAAAATGAGGCGGTGAATATCCCT 3136
 991 MetGlyMetGlyAlaAsnProTyrrGlyIleAlaIleAlaSerAsnGlnLe 1007
 |||||
 3137 ATGGGAATGGGGGCTAATCCCTATATGGCCAGACAGCATTAACCAACT 3186
 1007 uGlySerTrpProAspGlyMetLeuSerMetGlyGlnIleValSerHisGlyT 1024
 |||||
 3187 GGGTTCTGGCCCGATGATGCTTCCATGAGAACAGTTCTCATGGCA 3236
 1024 hArgIleAsnArgProLeuLeuArgAsnSerLeuAspAspLeuValGlyPro 1040
 |||||
 3237 CTCAAAATAGCCCTCTTCTAGGAATTCCTCGATGATCTTGTGTGGGCCA 3286
 1041 ProSerAsnLeuGlnGlnIleSerAspGlnArgAlaLeuLeuAspGlnLe 1057
 |||||
 3287 CTTTCCAACTCGGAAGCGCAGAGTGAAGAAAGACATTATTTGACCACT 3336
 1057 uHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuGlnGlnIleAspA 1074
 |||||
 3337 GCACAGCTCTTCTCAGCAACAGATGCCACAGGCGCTGGAAGAAATTTGACA 3386

1074 rGAlaLeuGlyIleProGluLeuValAsnGlnGlnAlaLeuGluPro 1090
 |||||
 3387 GAGCTTTGGCATTCCTGAATCTGTCAATCAAGGACAGCATTTAGAGCC 3436
 1091 LysGlnAspAlaPheGlnGlnGlnIleValAlaIleAlaValMetAspGlnLy 1107
 |||||
 3437 AAACAGATGCTTTCCAAAGCCCAAGAACACAGCATGAATGATGATCAGAA 3486
 1107 sAlaGlyLeuTyrrGlyGlnThrTyrrProAlaGlnGlyProProMetGlnG 1124
 |||||
 3487 GGCAGGATATATGACAGACATACCCAGCAGAGGGCTCCAAATGCAAG 3536
 1124 LysGlyPheHisLeuGlnGlnIleSerProSerPheAsnSerMetMetAsn 1140
 |||||
 3537 GAGGCTTTCATCTTCAGGAGCAATACCATCTTTAATCTATATGAT 3586
 1141 GlnMetAsnGlnGlnGlnLysAsnPheProLeuGlnGlnMetHisProArgAl 1157
 |||||
 3587 CAGATGAACCAAGCAAGCAATTTCTCTCTCCAGGAATGCACCCAGCAGC 3636
 1157 aAsnIleMetArgProArgThrAsnThrProLysGlnLeuArgMetGlnL 1174
 |||||
 3637 CAACATCATGAGACCCCGCAAAACCCCAACCACTTGAATGACAGC 3686
 1174 euGlnGlnArgLeuGlnGlnGlnIlePheLeuAsnGlnIleSerArgGlnAla 1190
 |||||
 3687 TTCAGCAGAGCTGCAGGGCCAGAGTTTGAATCAAGCCGACAGCA 3736
 1191 LeuGlnLeuLysMetGlyLysAsnProThrAlaGlyAlaIleAlaValMetAr 1207
 |||||
 3737 CTTGATTTGAATAATGGAATAACCTACTACTCTGCTGCTGGGCGATGAG 3786
 1207 gProMetMetGlnProGlnGlnGlyPheLeuAsnAlaGlnMetValAlaG 1224
 |||||
 3787 GCTTATGATGACGCCCGCAGAGGTTTCTTAATGCTCAATATGCTGCC 3836
 1224 lnaArgSerArgGluLeuLeuSerHisHisPheArgGlnGlnArgValAla 1240
 |||||
 3837 AACGCAGCAGAGACTGCTAAGTCATCACTTCCAGCAACAGAGGCTGCT 3886
 1241 MetMetMet.....GlnGlnGlnGlnGlnGlnGlnGlnGln 1253
 |||||
 3887 ATGATGATGACGACGACGACGACGACGACGACGACGACGACGACGACG 3936
 1253 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1270
 |||||
 3937 GCAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCA 3986
 1270 lAphSerProProProAsnValThrAlaSerProSerMetAspGlyLeu 1286
 |||||
 3987 CTTTACGCCACCTCTAATGTGACTGCTTCCCCAGCATGATGGGCTT 4036
 1287 LeuAlaGlyProThrMetProGlnAlaProProGlnGlnPheProTyrrG 1303
 |||||
 4037 TTGGCAGAGACCCACATATCCCAAGCTCTCCGCAACAGTTTCCATATCA 4086
 1303 nProAsnTyrrGlyMetGlyGlnGlnProAspProAlaPheGlyArgValS 1320
 |||||
 4087 ACCAAATTAATGGAATGGGACAACAACAGATCCAGCTTGTGCGAGTGT 4136
 1320 eRserProProAsnAlaMetMetSerSerArgMetGlyProSerGlnAsn 1336
 |||||
 4137 CTAGTCTCCCAATGCAATGATGTGTCMAAAGATGGGCTCTCCACAGAT 4186
 1337 ProMetMetGlnHisProGlnAlaIleSerIleTyrrGlnSerSerGlyMe 1353
 |||||
 4187 CCCATGATGACCAACCCGACGGCTGCATCATCTATAGTCTTAGAAT 4236
 1353 tLysGlyTrpProSerGlyAsnLeuAlaArgAsnSerSerPheSerGlnG 1370
 |||||
 4237 GAAGGGCGGCGCATCAGGAATTTGGCAGGAAACAGCTCTTTTCCAGC 4286
 1370 lGlnIlePheAlaHisGlnGlyAsnProAlaValTyrrSerMetValHisMet 1386


```
|||||
4287 ACCAGTTTCCCAACGAGGGAATCTGTCAGTGTATGATGGCCACATG 4336
1387 AaaglySerSerGlyHisMetGlyGlnMetAsnMetAsnProMetProke 1403
4337 AATGGCAGCAGTGTGTACATGGCAGATGATGAACATCAACCCCATGCCAT 4386
1403 tSerGlyMetProMetGlyProAspGlnLysTyrCys 1415
4387 GTCTGGCATGCTTATGGTCTGATCAGAAATATCTGTC 4423
```

seq_name: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:US-09-716-475-7321

```
seq_documentation_block:
: Sequence 7321, Application US/09716475
: GENERAL INFORMATION:
: APPLICANT: Shyjan, Andrew W.
: APPLICANT: Macbeth, Kyle J.
: APPLICANT: Vasicek, Thomas
: TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
: TITLE OF INVENTION: THEREFOR
: FILE REFERENCE: 1600.2037-001
: CURRENT APPLICATION NUMBER: US/09/716.475
: PRIORITY FILING DATE: 2000-11-20
: PRIOR APPLICATION NUMBER: 60/166,506
: NUMBER OF SEQ ID NOS: 8194
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7321
: LENGTH: 7888
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-716-475-7321
```

```
alignment_scores:
Quality: 7233.50      Length: 1429
Ratio: 5.174          Gaps: 3
Percent Similarity: 97.831      Percent Identity: 97.621
```

alignment_block:

US-09-041-994-2 x US-09-716-475-7321 ..

Align seg 1/1 to: US-09-716-475-7321 from: 1 to: 7888

```
1 MetSerGlyLeuGlyGlnAsnLeuAspProLeuAlaSerAspSerArgly 17
|||||
182 ATGAGTGTGATGTAGAGAAACTTGATCCACTGGCCAGTGCATTCACGAAA 231
17 sarGlyLysLeuProCysAspThrProGlyGlnGlyLeuThrCysSerGly 34
|||||
232 ACGCAAAATTCGATGTGATCTCCAGACAGGCTTACCTGCAGGTG 281
34 LuLysArgArgArgGlnGlnLysSerLysTyrLleGlnGlnLeuAlaGlu 50
|||||
282 AAAAAGCGAGCGGAGCAGAAAGTAAATATATTAAGAATGGCTGAG 331
51 LeuLleSerAlaAsnLeuSerAspLleAspAsnPheAsnValLysProAs 67
|||||
332 CTGATATCTGCCAATCTTAGTGATATTGACAATTTCAATGTCAAAACAGA 381
67 pLysCysAlaLleLeuLysGluThrValArgGlnLleArgGlnLleLysG 84
|||||
382 TAAATGTGGGATTTTAAAGAAACGTAAAGACAGATACGTCAAAATAAAG 431
84 LuGlnGlyLysThrLleSerAsnAspAspValGlnLysAlaAspVal 100
|||||
432 ACCAAGAAACTATTTCCATGTGATGATGATTTCAAAAAGCCGATGTA 481
101 SerSerThrGlyGlnGlyValLleAspLysAspSerLeuGlyProLeuLe 117
|||||
482 TCTTCTACAGGCGCAGGAGCTTATTTGATTAAGACTCTTAGAACCGCTTTT 531
```

```
117 uLeuGlnAlaLleuAspGlyPheLeuPheValValAsnArgGluAlaAsn 134
|||||
532 ACTTCAGGCATTTGATGATGTTTCTATTGTGGTAAATCGAAGCGAAACA 581
134 LeValPheValSerGluAsnValThrGlnTyrLleGlnLntyrLysGln 150
|||||
582 TTGTATTTGTATCAGAAATGTCAACAATACCTGCATTAATTAAGCAAGAG 631
151 AspLeuValAsnThrSerValTyrAsnLleLeuHisGlnLysPargLys 167
|||||
632 GACCTGGTTAAACACAGTGTATTTCATATCTTACATGAAGAACACGAAA 681
167 sAspPheLeuLysAsnLeuProLysSerThrValAsnGlyValSerTyrP 184
|||||
682 GCATTTTCTTAAGAAATTTACAAAATCTACAGTTAATGAGATTTCCTGGA 731
184 hrAsnGluProGlnArgGlnLysSerHisThrPheAsnGlySerMetLeu 200
|||||
732 CAATGAGACCCAAAGACAAAAGCCCATTCATTAATGCGCTATGTTG 781
201 MetLysThrProHisAspLleLeuGlnAspLleAsnAlaSerProGluMe 217
|||||
782 ATGAAAACACCATGATGATTTCTGGAAGACATTAAGCCAGTCTCAAT 831
217 LArgGlnArgTyrGluThrMetGlnCysPheAlaLeuSerGlnProArg 234
|||||
832 GCCCAGAGATATGAAACATGACAGTCTTGCCCTGCTCCAGCCAGAG 881
234 LMetMetGlnGlnGlnGlyAspLeuGlnSerCysMetLleCysValAla 250
|||||
882 CTTATGATGAGAGAGAGGAGGAGATTTTCAATCTTGATGATCTGTGGCA 931
251 ArgArgLleThrThrGlyGlnArgThrPheProSerAsnProGluSerPh 267
|||||
932 CGCCGATTAATCTACAGAGAGAAAGACATTTCCATCAACCTGAGAGCTT 981
267 eLleThrArgHisAspLeuSerGlyLysValValAsnLleAspThrAsn 284
|||||
982 TATTTACACACATGATCTTTCAGAAAGCTTGCATTAATATACAAAT 1031
284 eLleuArgSerSerMetArgProGlyPheGluAspLleLleArgArgCys 300
|||||
1032 CACTGAGATCTTCATGAGAGCGCTTGGAAGATTAATCCGAAAGCTT 1081
301 LLeGlnArgPhePheSerLeuAsnAspGlyGlnSerTyrSerGlnLysAr 317
|||||
1082 ATTCAAGATTTTATGCTTAATGATGGCAGTATGTTCCCAAGAAAG 1131
317 gHisTyrGlnGlu.....AlaTyrL 324
|||||
1132 TCACATATCAAGAAAGTTTACCAAGTATGATGATTTTCCCAACAGCTTATC 1181
324 euAsnGlyHisAlaGluThrProValTyrArgPheSerLeuAlaAspGly 340
|||||
1182 TTAATGCGCATGACAGAAACCCAGTATATCGATTTCTGTGGCTGATGGA 1231
341 ThrLleValThrAlaGlnThrLysSerLysLeuPheArgAsnProValTh 357
|||||
1232 ACTATAGTGACAGCAGACAAAAGCAAACTTCCGAATTCCTGTATAC 1281
357 rAsnAspArgHisGlyPheValSerThrHisPheLeuGlnArgGlnGln 374
|||||
1282 AATATGATGACATGCGCTTGTCTCAACCCACTTCCAGAGAAACAGA 1331
374 snGlyTyrArgProAsnProAsnProValGlyGlnGlyLleArgPropo 390
|||||
1332 ATGGAATATGACCAACCAATCCGTGTGGACAAAGGATTAAGACCACT 1381
391 MetAlaGlyCysAsnSerSerValGlyGlyMetSerMetSerProAsnG 407
|||||
1382 ATGGCTGATGCAACAGTTCGGTAGGCGGCATGATGATTCGCCAACA 1431
407 nGlyLeuGlnMetProSerSerArgAlaTyrGlyLeuAlaAspProSerT 424
```



```

|||||
1432 AGGCTTACAGATCCGACGACGAGGCGCTATGCGCTTGGCAGACCTTAGCA 1481
424 hTThGClYlInMetSerGlyAlaArgTyrGlySerSerAsnIleAla 440
1482 CCACAGGGGAGATGAGTGGAGTAGTATGGGGTTCAGTAACATAGCT 1531
441 SerLeuThProGlyProGlyMetGlnSerProSerSerTyrGlnAsnAs 457
1532 TCATTGACCCCTGGGCGGACGATGCATACCATCTTCTTACCAGAACAA 1581
457 nAsnTyrGlyLeuAsnMetSerSerProProHisGlySerProGlyLeuA 474
1582 CAACATFAGGCTCAACATGATGATAGCCCCACACATGGAGAGTCCCTGG 1631
474 lAProAsnGlnGlnAsnIleMetIleSerProArgAsnArgGlySerPro 490
1632 CCCCMAACGACGAGATATCATGATTTCTCTGTAATCGTGGAGTCCA 1681
491 LysIleAlaSerHisGlnPheSerProValAlaGlyValHisSerProke 507
1682 MAGATAGCTTCATCATGATTTCTCTGTTGCAAGGTGTCACACTCTCCAT 1731
507 lAlaSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerSerLeus 524
1732 GGCATCTTCTGGCAATACTGGGACACAGACCTTTCACAGACTCTCTCA 1781
524 eAlAlaLeuGlnAlaIleSerGlyGlyValGlyThrSerLeuSerThr 540
1782 GTGCCCTTGACACCATCATGAGTGAAGGTGTGGGACCTTCTTATCTACT 1831
541 LeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIlePh 557
1832 CTGTCTATCCAGGCCCCCAATTTGATTACTCTCCCAATATGAAATTAAC 1881
557 rGlnProSerLysValSerAsnGlnAspSerLysSerProLeuGlyPheT 574
1882 CCAACCAAGTAAAGTAAAGCAATCAGAGATCCAAAGAGTCCCTGGGGTTT 1931
574 yTcYAspGlnAsnProValGlySerSerMetCysGlnSerAsnSerArg 590
1932 ATTTGGACCAAAATCCAGTGAAGAGTTCAATGTTCAGTCAAAATAGCAGA 1981
591 AspHisLeuSerAspLysGlySerLysGlySerSerValGlyValAlaG 607
1982 GATCACCTTAGTGACAAAGAAAGTAAAGAGAGAGTGTGAGGGGGCAGA 2031
607 uAsnGlnArgGlyProLeuGlySerLysGlyHisLysLysLeuLeuGln 624
2032 GAATCAAAAGGGTCCCTTGGAAAGCAAGGTCAATAAAATTAATGCACT 2081
624 euLeuThrCysSerSerAspAspArgGlyHisSerSerLeuThrAsnSer 640
2082 TACTTACCTGTTCTCTGATGACCGGGGTCAATTCCTTGTGACCAACTCC 2131
641 ProLeuAspSerSerCysLysGlySerSerValSerValThrSerProse 657
2132 CCCCTAGATTCAAGTTGTAAGAAATCTCTGTAGTGCACAGCCCCCTC 2181
657 rGlyValSerSerSerThrSerGlyValSerSerThrSerAsnMetH 674
2182 TGGAGCTCTCTCTCATCATCTGAGAGAGTATCCCTCATCATCAATATGC 2231
674 lAsGlySerLeuLeuGlnGlyLysHisArgIleLeuHisLysLeuLeuGln 690
2232 ATGGGCTACTGTTTACAAAGAGACACCGGATTTTGCACAAGTTTGCAGC 2281
691 AsnGlyAsnSerProAlaGlyValAlaLysIleThrAlaGlnAlaThrG 707
2282 AATGGGAATTTCAACAGCTAGGTAGCCAAAGATTTACTGCAGAAAGCCCTG 2331
707 yLysAspThrSerSerIleThrSerCysGlyAspGlyAsnValValLysG 724
|||||

```

```

2332 GAAAGACACCAGCAGTATAACTTCTTGGGGACGGAAATGTTGTCAACG 2381
724 InGlnGlnLeuSerProLysLysGlyLysAsnAsnAlaLeuLeuArgTyr 740
2382 AGAGCAGCTAAGTCTTAAGAAAGAGAAATATGCACTTCTTAATATAC 2431
741 LeuLeuAspArgAspAspProSerAspAlaLeuSerLysGlyLeuGlnPr 757
2432 CTGTCTGACAGGAGATATCTTAAGTATGCACTCTTAAGAAACTACAGCC 2481
757 oGlnValGlyGlyValAspAsnLysMetSerGlnCysrThrSerSerThr 774
2482 CCAAGTGAAGAGAGTGGATTAATTAATGATGATGATGATGATGATGATG 2531
774 lAProSerSerSerGlnGlnLysAspProLysIleLysThrGlyThrSer 790
2532 TTCTTAAGCTCAAGTCAAGAAAGACCTTAATTAATTAAGACAGACAACT 2581
791 GluGlyGlySerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLe 807
2582 GAAAGAGGATCTTGAGACTTGGATTAATCTAGATGCTATTTCTGTGATCT 2631
807 uThrSerSerAspPheTyrAsnAsnSerIleSerSerAsnGlySerHisL 824
2632 GACTACTTCTGACTTTTACATATTAATTCATATCTTCAATGATGATGATC 2681
824 euGlyThrLysGlnGlnValPheGlnGlyThrAsnSerLeuGlyLeuLys 840
2682 TGGGACTTAAGCAACAGGTGTTCAAAGGAACTAATCTCTGGGTTTGAAA 2731
841 SerSerGlnSerValGlnSerIleArgProProTyrAsnArgAlaValSe 857
2732 AGTTTACAGTCTGTGAGTCTATCTCTCTCCATATATACCGAGCACTGTC 2781
857 rLeuAspSerProValSerValGlySerSerProProValLysAsnIleS 874
2782 TCTGATATCCCTGTTTCTGTGGCTCAAGTCTCCAGTAATAAAATATACA 2831
874 eAlAlaPheProMetLeuProLysGlnProMetLeuGlyGlyLysnProArg 890
2832 GTGCTTTCCTCCATGTTTACCAAAAGCAACCATGTTGGGTGGAATCCAAA 2881
891 MetMetAspSerGlnGlnAsnTyrGlySerSerMetGlyGlyProAsnAr 907
2882 ATGATGATATGTCAGGAATAATTTGCTCAAGTATG..... 2917
907 gAsnValThrValThrGlnThrProSerSerGlyAspTrpGlyLeuProA 924
2918 .....GGAGACTGGGGCTTATCCAA 2936
924 snSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyArg 940
2937 ACTCAAAAGCCCGGACAAATGAACTATGAACTCAATCAACTCATGGGAAA 2986
941 ProGlyGlyAspPheTyrAsnThrSerLeuProArgProAlaLeuGlyLys 957
2987 CAGGAGGAGATTATATATCTTCTTACCCAGACGCTGAGTGGGTGCTC 3036
957 rIleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgProV 974
3037 TATTTCCACATTTGCTCTCTCGGTCTAATATAGCATACAGGTGGAGACAG 3086
974 aLleuGlnGlnGlnGlnGlnMetLeuGlnMetArgProGlyGlyLysIlePro 990
3087 TATTTGCAACAGCAGCAGCAGATGCTTCAATATAGGCTGTGTAATCCCC 3136
991 MetGlyMetGlyAlaAsnProTyrGlyGlnAlaAlaAlaSerAsnGlnLe 1007
3137 ATGGGAATGGGGCTATCTCTATGSCCAAGCAGCAGCTATACCAACT 3186
1007 uGlySerTrpProAspGlyMetLeuSerMetGlnGlnValSerHisGlyT 1024
3187 GGGTTCCTGGCCGATGATGATGATGATGATGATGATGATGATGATGATG 3236

```

1024 hrglnasnArgProLeuLeuArgAsnSerLeuAspLeuValGlyPro 1040
|||||
3237 CTCAAATAGGCCCTCTTAGGAATCCCGATGATCTGTGGGCCA 3286
1041 ProSerAsnLeuGlnGlnSerAspGluArgAlaLeuLeuAspGlnLe 1057
|||||
3287 CCTCCCAACCTGGAAAGCCAGAGTGAAGAAAGACATTATTTGGCCAGCT 3336
1057 uHisThrLeuSerAsnThrAspAlaThrGlyLeuGluGluLeuAsp 1074
|||||
3337 GCACACTCTTCTCAGCAACACAGATGCCACAGCCTGGAGAAATTGACA 3386
1074 rgaAlaLeuGlyIleProGluLeuValAsnGlnGlnAlaLeuGluPro 1090
|||||
3387 GAGCTTTGGGCATTCTGAACCTGTCAATCAGGACAGGCATTAGAGCCC 3436
1091 LysGlnAspAlaPheGlnGlnGlnGluAlaAlaValMetAspGlnIly 1107
|||||
3437 AACACAGATGCTTCCAAAGGCCAAGAACAGCACTAATGATGATCAGAA 3486
1107 sAlaGlyLeuTyrGlyGlnThrTyrProAlaGlnGlyProPrometGln 1124
|||||
3487 GGCAGGATTTATGACACAGACATCCACAGCAGGGCCTCCAAATGCAG 3536
1124 LysGlyPheHisLeuGlnGlyGlnSerProSerPheAsnSerMetLeuAsn 1140
|||||
3537 GAGGCTTTCATCTTCAGGACAAATCACCATCTTTAACTATGATGAAT 3586
1141 GlnMetAsnGlnGlnGlnGlnAsnPheProLeuGlnGlyMetHisProArgAl 1157
|||||
3587 CAGATTGAACCGCAGAGCAATTTCTCTCCAAAGAAATGCACCCAGCAG 3636
1157 aAsnIleMetArgProArgThrAsnThrProLysGlnLeuArgMetGlnI 1174
|||||
3637 CAACATCATGAGACCCCGGACAAACACCCCAACCACTTAGAATGCAGC 3686
1174 euGlnGlnArgLeuGlnGlnGlnGlnPheLeuAsnGlnSerArgGlnAla 1190
|||||
3687 TTCAGCAGAGGCTCAGGGCCAGCAGTTTGAATCAGAGCCGACAGGCA 3736
1191 LeuGlnLeuLeuMetGluAsnProThrAlaGlyIlyAlaAlaValMetArg 1207
|||||
3737 CTTGAATTGAAATGAAACCTACTAGCTGTGTGCTGCTGCGGTGATGAG 3786
1207 gPrometMetGlnProGlnGlnGlyPheLeuAsnAlaGlnMetValAlaG 1224
|||||
3787 GCCATATGATGAGCCCGCAGCGGTTTCTTAATGCTCAATATGCTCGCCC 3836
1224 LnaArgSerArgGluLeuLeuSerHisHisPheArgGlnGlnArgValAla 1240
|||||
3837 AACGCACAGAGAGCTGTAAGTCATCACTCCGACAAACAGAGGTGGCT 3886
1241 MetMetMet.....GlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1253
|||||
3887 ATGATGTATGACGACGACGACGACGACGACGACGACGACGACGACGAC 3936
1253 nglnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1270
|||||
3937 GCAGCAGACAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3986
1270 LapeSerProProProAsnValThrAlaSerProSerMetAspGlyLeu 1286
|||||
3987 CCTTCACGCCCACTCCTAAATGTGACTGCTTCCCGACATGATGAGGCTT 4036
1287 LeuAlaGlyProThrMetProGlnAlaProProGlnGlnPheProTyrGly 1303
|||||
4037 TTGGCAGGACCCCAATGCCCAGAGCTCTCCGCAACAGTTTCCATATCA 4086
1303 nProAsnTyrGlyMetGlyGlnGlnGlnProAspProAlaPheGlyArgValS 1320
|||||
4087 ACCAAATTATGGAATGGACACAACAGATCCAGCCTTTGGTCGAGTGT 4136

1320 erSerProProAsnAlaMetMetSerSerArgMetGlyProSerGlnAsn 1336
|||||
4137 CTAGTCTCCCAATGCAATGATGTCTCAAGAAATGGCTCCCTCCAGAAAT 4186
1337 ProMetMetGlnHisProGlnAlaAlaSerIleTyrGlnSerSerGluMe 1353
|||||
4187 CCCATGATGCAACACCCGACAGGCTGCATCCATCTATAGTCTCAGAAAT 4236
1353 LysGlyTyrProSerGlyAsnLeuAlaArgAsnSerSerPheSerGlnG 1370
|||||
4237 GAAGGCTGGGCATCAGGAATTTGGCCAGAGAACAGCTCTTTTCCAGC 4286
1370 LnglnPheAlaHisGlnGlyAsnProAlaValTyrSerMetValHisMet 1386
|||||
4287 AGCAGTTTGGCCACAGGGAATCTCAGTGTATAGTATGATGTCACATG 4336
1387 AsnGlySerSerGlyHisMetGlyGlnMetAsnMetAsnPrometProm 1403
|||||
4337 AATGGCAGCAAGTGTCAATGGACAGATGAACATGAAACCCCATGCCAT 4386
1403 tSerGlyMetProMetGlyProAspGlnLysTyrCys 1415
|||||
4387 GTCTGGCATGCTATGGGTCTCTGATCAGAAATACTGC 4423

This Page Blank (uspto)

OM of: US-09-041-994-2 to: Pending_Patents_NA.New:* out_format : pfs
Date: May 1, 2001 6:24 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=framed.p2n.model -DEV=xlp  
-Q=/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-196-296B-1 + 7341.00 6467.41 0.0  
-DB=Pending_Patents_NA.New -OFMT=fastap -SUFFIX=p2n.rnpn  
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000  
-LOOPEXT=0.000 -OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000  
-XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -XGAPOP=10.000  
-XGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200  
-THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=2000000000  
-USER=US09041994_@CGN1_1.175 -NCPV=6 -ICPV=3 -LONGLOG -NO_XLPHY  
-MAIT -THREADS=1
```

Search information block:

```
Query: US-09-041-994-2  
Query length: 1415  
Database: Pending_Patents_NA.New:*  
Database sequences: 201016  
Database length: 146721182  
Search time (sec): 114.180000
```

Score list:

Sequence	Strd Orig	ZScore	Escore	len	Documentation
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-196-296B-1 + 7341.00 6467.41 0.0				47	
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-196-296B-3 + 480.00 421.69 4.56-16				33	
/cg2_6/ptodata/2/pna/PCT_NEW_COMB.seq:PCT-US01-01339-984A - 299.50 247.56 2.2e-06					
/cg2_6/ptodata/2/pna/PCT_NEW_COMB.seq:PCT-US01-01339-7281 + 226.00 187.65 0.0048					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12207 - 226.00 151.43 0.0502					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12216 - 224.00 153.49 0.3879					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12217 + 218.50 149.88 0.6159					
/cg2_6/ptodata/2/pna/US08_NEW_COMB.seq:US-09-956-171C-59 + 216.00 172.11 0.0356					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12208 - 216.00 150.03 0.6043					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-654-935A-142 + 215.50 187.81 1.19					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12218 + 215.50 144.74 0.148					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-813-206-703 - 210.50 178.94 0.0148					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-543-679A-3003 - 210.00 165.83 0.0797					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12214 - 207.50 141.42 1.82					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12216 + 207.50 138.94 2.51					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-543-679A-2410 - 205.50 175.03 0.0245					
/cg2_6/ptodata/2/pna/US06_NEW_COMB.seq:US-09-248-505-604 + 205.50 160.00 0.1682					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12205 - 205.00 137.71 2.93					
/cg2_6/ptodata/2/pna/US06_NEW_COMB.seq:US-09-248-823-10 - 204.00 158.46 0.2050					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-813-925-1 + 203.00 171.24 0.0398					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12211 - 202.50 138.29 2.72					
/cg2_6/ptodata/2/pna/US06_NEW_COMB.seq:US-09-254-168-41 + 199.50 155.42 0.3029					
/cg2_6/ptodata/2/pna/US06_NEW_COMB.seq:US-09-801-833-7501 + 191.50 167.25 0.0664					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-792-024-57 + 191.00 169.00 0.0530					
/cg2_6/ptodata/2/pna/US06_NEW_COMB.seq:US-09-248-505-428 + 191.00 153.93 0.3664					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-649-996-5 + 190.50 159.03 0.1905					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12218 - 190.00 122.26 21.29					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-739-449-211 - 187.50 130.89 7.04					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12215 - 187.00 119.71 29.54					
/cg2_6/ptodata/2/pna/US06_NEW_COMB.seq:US-09-281-593-237 - 186.50 157.68 0.2265					
/cg2_6/ptodata/2/pna/US08_NEW_COMB.seq:US-09-956-171C-264 - 184.00 157.48 0.2324					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-543-679A-2717 + 184.00 153.60 0.3822					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-543-679A-2752 + 184.00 148.44 0.7409					
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12211 + 183.00 121.10 24.69					

```
/cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12215 + 182.50 115.74 49.13  
/cg2_6/ptodata/2/pna/PCT_NEW_COMB.seq:PCT-US01-01339-5749 + 181.00 156.98 0.2477  
/cg2_6/ptodata/2/pna/PCT_NEW_COMB.seq:PCT-US01-01339-9949 - 181.00 156.98 0.2477  
/cg2_6/ptodata/2/pna/PCT_NEW_COMB.seq:PCT-US01-08117-122 + 181.00 153.74 0.3757  
/cg2_6/ptodata/2/pna/US06_NEW_COMB.seq:US-09-248-505-39 + 181.00 143.25 1.44
```

```
seq_name: /cg2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-196-296B-1
```

seq_documentation_block:

```
; Sequence 1, Application US/09196296B  
; GENERAL INFORMATION:  
; APPLICANT: Suen, Chen-Shian  
; APPLICANT: Frail, Donald E.  
; APPLICANT: Lytle, Richard C.  
; TITLE OF INVENTION: Cloning and Expression of a Nuclear  
; FILE REFERENCE: 0630/01376  
; CURRENT APPLICATION NUMBER: US/09/196,296B  
; CURRENT FILING DATE: 1998-11-19  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 4789  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-09-196-296B-1
```

alignment_scores:

```
Quality: 7341.00 Length: 1423  
Ratio: 5.192 Gaps: 2  
Percent Similarity: 99.368 Percent Identity: 99.157
```

```
alignment_block:  
US-09-041-994-2 x US-09-196-296B-1 ..
```

```
Align seg 1/1 to: US-09-196-296B-1 from: 1 to: 4789
```

```
1 MetterGlyLeuGlyGlyLeuAspProLeuAlaSerAspArgly 17  
|||||  
185 ATGAGTGGATTGAGGAGAACTTGATCCAGTCGACGATGTTACGAAA 234  
17 sarGlyLeuProCysAspThrProGlyGlyLeuThrCysSerGly 34  
|||||  
235 ACGGAATGCCATGCTATCTATCTCCAGGACGATCTTACCTGCGTGTG 284  
34 IulysArgArgArgGlyGlyGlyLeuSerIlyrIleGlyGlyLeuAlaGly 50  
|||||  
285 AAAACGGAGACGAGGAGGAGAAAGTAATTAATTTGAAGATTCGTCGAG 334  
51 IeulIeSerAlaAsnLeuSerAspIleAspAsnPhesAsnValIysProAs 67  
|||||  
335 CTGATATCTGCCAATCTTAGATATTCACAAATTCATGCAATGCAACACGA 384  
67 PLYCysAlaIleLeuYsgIuThrValArgGlnIleArgGlnIleYsg 84  
|||||  
385 TAAATGCGCATTTTAAAGCAACAGTAGACATGCGTCAAAATAAAG 434  
84 IuGlnGlyIlyThrIleSerAsnAspAspValGlnIlyAlaAspVal 100  
|||||  
435 AGCAAGGAAACATTTTCCATATGATATGTTCAAAAAACCGATGTA 484  
101 SerSerThrGlyGlyIlyValIleAspIlyAspSerLeuGlyProLeu 117  
|||||  
485 TCTTCTACAGGCGGAGGAGTATGATTAAGACCTTTAGACCGCTTTT 534  
117 IuGlnGlnAlaLeuAspIlyPheLeuPheValIleAsnArgGlnAlaAsn 134  
|||||  
535 ACTTCAGCATTTGATGATGTTCTATTGTGGTGAATGAGGAGGAACA 584  
134 IeValPheValSerGluAsnValThrGlnIlyLeuGlnIlyGlnGln 150  
|||||  
585 TTGATTTTGTATCGAAATGTCACAAATACCTGCATTAATTAACGAAG 634
```

151 AspleuValAsnThrSerValTyrAsnIleuHisGluAspArgLy 167
 |||
 635 GACCTGGTTAACACAGTGTACATATCTACATGAAAGACGAA 684
 167 sAspHeuLysAsnLeuProLysSerThrValAsnGlyValSerTrp 184
 |||
 685 GGAATTTCTTAGAATTTACCAAAATCTACAGTTAATGAGTTTCTGGA 734
 184 hAsnGluProGlnArgGlnLysSerHisThrPheAsnCysArgMetLeu 200
 |||
 735 CAATGAGACCAAGACAAAAAGCCATACATTTAATTGCCGTATGTTG 784
 201 MetLysThrProHisAspIleLeuGluAspIleAsnAlaSerProGlu 217
 |||
 785 ATGAAACACCAACATGATATCTGGAAGACATAACGCCAGTCTGAAAT 834
 217 fArgGlnArgTyrGluThrMetGlnCysPheAlaLeuSerGlnProArg 234
 |||
 835 GGCACGAGATATGAACCAATGCAAGTGTGGCTGTCTCAGCCAG 884
 234 lMetMetGluGluGlyGluAspLeuGlnSerCysMetIleCysValAla 250
 |||
 885 CATATGATGAGAGAGGGAAGATTTCATCTTGATGATCTGTGTGCA 934
 251 ArgArgIleThrThrGlyGluArgThrPheProSerAsnProGluSerPh 267
 |||
 935 CCCCCATTACTACAGAGAAAGAACATTTCCATCAACCCCTGAGACT 984
 267 eIleThrArgHisAspLeuSerGlyLysValValAsnIleAspThrAsn 284
 |||
 985 TATTCACAGACATGATCTTTCAGAAAGGTGTCAATATAGATACAAAT 1034
 284 eLeuArgSerSerMetArgProGlyPheGluAspIleLeuArgCys 300
 |||
 1035 CACTGAGATCCCTCCATGAGGCGCTTGAGATATATCCGAAAGTGT 1084
 301 lIleGlnArgPhePheSerLeuAsnAspGlyGlnSerTrpSerGlnLysAr 317
 |||
 1085 ATTCAGAGATTTTTTGTAGCTAATGATGGGCAATGCTGCCAGAAAG 1134
 317 gHisTyrGlnGluAlaTyrLeuAsnGlnLysAlaGluThrProValTyr 334
 |||
 1135 TCATCTACAGAAAGCTTATCTTAATGCGCATGACAGAACCCCACTATATC 1184
 334 rGpHeSerLeuAlaAspGlyThrIleValThrAlaGlnThrLysSerLys 350
 |||
 1185 GATTTCTGTTGGCTGATGGAACTATGATGCTGCACAGACAAAAGCATA 1234
 351 LeuPheArgAsnProValThrAsnAspArgHisGlyPheValSerThrH 367
 |||
 1235 CTCTTCGGAATCTGTAACTAATGATGACATGGCTTTGTCTCAACCCA 1284
 367 sPheLeuGlnArgGluGlnAsnGlyTyrArgProAsnProAsnProValG 384
 |||
 1285 CTTCCTTCAGAGAGAACAGATGATATGACCAACCAATCTCTGTG 1334
 384 lYcGlnGlyIleArgProPrometAlaGlyCysAsnSerSerValGlyGly 400
 |||
 1335 GACAAAGGATTAGACCACTATGGCTGATGCAACAGTTGCGTAGCGCG 1384
 401 MetSerMetSerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTyr 417
 |||
 1385 ATGAGATATCTGCGCAAAACCAAGCTTACAGATGCCAGAGCGGCGCTA 1434
 417 rGlyLeuAlaAspProSerThrThrGlyGlnMetSerGlyAlaArgTyrG 434
 |||
 1435 TGGCTTGGAGACCTTAGCACCAACAGGCGAGATGATGAGCTAGGTATG 1484
 434 lYcLysSerSerAsnIleAlaSerLeuThrProGlyProGlyMetGlnSer 450
 |||
 1485 GGGGTTCCAGTAAACATAGCTTATGACCCGTGGCCAGGACATGCAATCA 1534

451 ProSerSerTyrGluAsnAsnAsnTyrGlyLeuAsnMetSerSerPro 467
 |||
 1535 CCATCTTCTTACCAGAACACAACTATGGCTCAATCAGTAGGCCCCC 1584
 467 oHisGlySerProGlyLeuAlaProAsnGlnGlnAsnIleMetLieserP 484
 |||
 1585 ACATGGGAGTCTCTGCTTGGCCCCAAACAGCAGAAATATCATGATTTTC 1634
 484 rArgAsnArgGlySerProLysIleAlaSerHisGlnPheSerProVal 500
 |||
 1635 CTGCTATGCTGGAGAGTCCAAAGATAGCTCATATCAGTTTCTCTGTT 1684
 501 AlaGlyValHisSerPrometAlaSerSerGlyAsnThrGlyAsnHisSe 517
 |||
 1685 GCAGGTGCACCTCTCCATGTCATCTTGTGCAATACGCGAACCCAG 1734
 517 rPheSerSerSerSerLeuSerAlaLeuGlnAlaLieserGlyValG 534
 |||
 1735 CTTTTCACAGACCTCTCTCAGTGCCTGCCAAGCATCATGAAAGTGTGG 1784
 534 lYThrSerLeuLeuSerThrLeuSerSerProGlyProLysLeuAspAsn 550
 |||
 1785 GGACTTCCCTTTATCTACTCTGTCAATCACAGGCCCAATTTGGATATC 1834
 551 SerProAsnMetAsnIleThrGlnProSerLysValSerAsnGlnAspSe 567
 |||
 1835 TCCTCCAAATATGATATATACCAACCAAGTAAAGTAAAGTACAGATTC 1884
 567 rLysSerProLeuGlyPheTyrCysAspGlnAsnProValGluSerSerM 584
 |||
 1885 CAAGAGTCTCTGCTGGCTTTATTTGCGACCAAAATCCAGTGGAGAGTTCA 1934
 584 eCysGlnSerAsnSerArgAspHisLeuSerAspLysGluSerLysGlu 600
 |||
 1935 TGTGTACGTAATATAGACAGATCACCTCAGTGACAAAGAAAGTAAAGAG 1984
 601 SerSerValGluGlyAlaGluAsnGlnArgGlyProLeuGluSerLysG 617
 |||
 1985 AGCAGGTGTAGAGGGGCAAGAAATCAAGGGGCTTTGGAAGCAAAAG 2034
 617 yHisLysLysLeuLeuGlnLeuLeuThrCysSerSerAspAspArgGlyH 634
 |||
 2035 TCATTAATAATATATCTGACATTAATCTTCTTGTGATGACGGGGGTC 2084
 634 lAsSerSerLeuThrAsnSerProLeuAspSerSerCysLysGluSerSer 650
 |||
 2085 ATTCCTCTTGACCAACTCCCTCTGATTAAGTGTAAAGAACTGTCT 2134
 651 ValSerValThrSerProSerGlyValSerSerSerThrSerGlyGly 667
 |||
 2135 GTTAGTGTACCAAGCCCTCTGGAGTCTCTCTCTTACATCTGGAGAGT 2184
 667 lSerSerThrSerAsnMetHisGlySerLeuLeuGlnGlyLysHisArgI 684
 |||
 2185 ATCTCTTACATCCAAATATGATGATGCTGACTGTACAAAGAAAGCACCGGA 2234
 684 lLeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAlaLys 700
 |||
 2235 TTTTGCACAAGTTGCTGCAGAAATGGGAATTCACCACTGAGGTAGCCAAG 2284
 701 lIleThrAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysG 717
 |||
 2285 ATTAGTGCAGAAAGCCACTGGGAAAGACACACAGATATACATCTTGTGG 2334
 717 yAspGlyAsnValValLysGlnGlnLeuLeuSerProLysLysGlyAla 734
 |||
 2335 GGACGGAATGTGTCTCAACAGAGAGCAGCTAAGTCTTAAGAAAGAGAGA 2384
 734 snAsnAlaLeuLeuArgTyrLeuLeuAspArgAspAspProSerAspAla 750
 |||
 2385 ATATATCAGCTTCTTAGATACCTGCTGACAGGAGATGATCTTGTATGCA 2434
 751 LeuSerLysGluLeuGlnProGlnValGluGlyValAspAsnLysMetSe 767

```

|||||
2435 CTCTCTAAGACAGACAGCCCAAGTGGAGAGATGATATAATAAGAG 2484
767 rGlnCysThrSerSerThrIleProSerSerSerGlnGluLysAspProL 784
|||||
2485 TCAGTGCACACAGCTCCACCATCTTCCATCAGAGTCAAGAGAACCCCTA 2534
784 ysiIeuysThrGluThrSerGlnGluGlySerGlyAspLeuAspAsnLeu 800
|||||
2535 AAATTAAAGACAGACAGAGTGAAGAGGATCTGGAGACTTGGAATTAATCTA 2584
801 AspaIaIleuGlyAspLeuThrSerSerAspPheTyrAsnAsnSerIle 817
|||||
2585 GATGCTATCTTGCTGTGATCTGACAGTTCGACTTTACATAATATTCAT 2634
817 eSerSerAsnGlySerHisIeuGlyThrLysGlnIValPheGlnGlyT 834
|||||
2635 ATCCTCAAAATGCTAGTCATCTGGGAGCTAAAGCAACAGGTGTTCAAGAA 2684
834 hirsAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleArgPro 850
|||||
2685 CTAAATCTCTGGGTTTGAAGTACACAGTCTGTGACGTCTATTGCTCT 2734
851 ProTyrAsnArgAlaValSerIeuAspSerProValSerValGlySerSe 867
|||||
2735 CCATATAACGAGCAGCTGCTCTGTGATAGCCCTTTCTGTGGCTCAG 2784
867 rProValIleAsnIleSerAlaPheProMetLeuProLysGlnProM 884
|||||
2785 TCCCGACGTAAATAATTCAGTCTTCCCATGTATCAAAAGCAACCCA 2834
884 etIeuGlyGlyAsnProArgMetMetAspSerGlnIleAsnTyrGlySer 900
|||||
2835 TGTTGGGTGGGAATCCAGATGATGATGTCAGAAATTAATGCTCA 2884
901 SerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProSerSe 917
|||||
2885 AGTATGGGTGGGCCAAACCCAAATGTGACTGTGCTCAGACTCCTTCTCT 2934
917 rGlyAspTyrGlyLeuProAsnSerLysAlaGlyArgMetGluProMetA 934
|||||
2935 AGGAGACTGGGGCTTACCAAACTCAAAAGCCGCGCAGATGGAACCTTA 2984
934 snSerAsnSerMetGlyArgProGlyGlyAspTyrAsnThrSerLeuPro 950
|||||
2985 ATTCAAACTCCATGGGAGAGACAGAGAGAGATTAATTAATCTTTTACCC 3034
951 ArgProAlaIeuGlyGlySerIleProThrIleuProLeuArgSerAsnSe 967
|||||
3035 AGACCTGCACCTGGGTGGCTCTATTCCACAGTTGGCTCTTCGCTTAATAG 3084
967 rIleProGlyAlaArgProValIeuGlnGlnGlnGlnMetLeuGlnM 984
|||||
3085 CATACCAAGTCCGAGACAGATTTGCAACACAGACAGACAGATGCTTCAA 3134
984 etArgProGlyGlyIleProMetGlyMetGlyAlaAsnProTyrGlyGln 1000
|||||
3135 TGAGGCTGTGGAAATCCCATGGGAATGGGGGCTTAATCCCTATGGCCAA 3184
1001 AlaIaIaIaSerAsnGlnIleuGlySerThrProAspGlyMetLeuSerMe 1017
|||||
3185 GCACGACATCTTAACCAACTGGGTTCCTGGGCCCATGGCATGTGTCCAT 3234
1017 tGluGlnIValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerL 1034
|||||
3235 GGAACAAGTTTCTCATGCGACTCAAAATAGGCCCTTCTTAGAATGCC 3284
1034 euAspAspLeuValGlyProProSerAsnIeuGlyGlnIleSerAspGlu 1050
|||||
3285 TGGATGTATCTTGTGGGCGACCTTCCAAACTGGAAGGCGACAGAGAGCAA 3334
1051 ArgAlaIeuLeuAspGlnIleuHisThrIleuSerAsnThrAspAlaTh 1067
|||||

3335 AGACATTATTGACAGCTGACACTCTTCTCAGCAACACAGATGGCAC 3384
1067 rGlyLeuGlnGluIleAspArgAlaIleuGlyIleProGluLeuValAsnG 1084
|||||
3385 AGGCTGTGAAGAAATTCAGACAGCTTGGCATTCCTGAACTGTCTCATC 3434
1084 IngIyGlnAlaIeuGluProLysGlnAspAlaPheGlnGlnGlnAlaI 1100
|||||
3435 AGGACACAGCATTTAGACCCCAACAGCATGCTTCCAAAGCCCAAGAAACA 3484
1101 AlaValMetMetAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAl 1117
|||||
3485 GCAGTAAATGATGATCAGAAAGCGAGATTATGAGACAGACATGCCAGC 3534
1117 aGlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlyGlnSerProS 1134
|||||
3535 ACAGGGGCTTCCAAATGCAGAGAGCTTTCATCTTCAGGAGCAATACCAT 3584
1134 erPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlnIleAsnPheProLeu 1150
|||||
3585 CTTTAACCTCATGATGATCAATCAGATGAACCAAGCAACATTTTCTCTC 3634
1151 GlnGlyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPr 1167
|||||
3635 CAAGGAATGCACCCACGACGCCAATCATGAGACCCCGGACAAACACCCC 3684
1167 OlySGlnLeuArgMetGlnIleuGlnGlnArgLeuGlnGlyGlnIlePheL 1184
|||||
3685 CAACCAACTTAGAATGCAGCTTCAGCAGAGGCTTCAGGGCCAGCAGTTTT 3734
1184 euAsnGlnSerArgGlnAlaIleuGlnIleuLysMetGluAsnProThrAla 1200
|||||
3735 TGAATCAGAGCCGACAGGCACTTGAATTAATGAATAAGAAACCCATGCT 3784
1201 GlyIyAlaIaIaValMetArgProMetMetGlnProGln.....G 1214
|||||
3785 GGTGTGCTGTGGATGATGAGGCTATGATGAGCCGACCGAGGAGCTGCCA 3834
1214 nGlyPheLeuAsnAlaGlnMetValAlaGlnArgSerArgGluLeuLeuS 1231
|||||
3835 GGGTTTTCTTAATGCTCAAAATGGTGCGCCCAACGACGACGAGAGCTGTA 3884
1231 erHisHisPheArgGlnGlnArgValAlaIleMetMetMet..... 1243
|||||
3885 GTCAATCACTTCCGACACAGAGGGGTGCTATGATGATGACAGACACCA 3934
1244 ...GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1259
|||||
3935 CAGCAGCAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3984
1259 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1276
|||||
3985 ACAGCAACAGCAGCAGCAGCAGCAGCAACCCAGCCTTCAGGCCACCTCT 4034
1276 snValThrAlaSerProSerMetAspGlyLeuLeuAlaGlyProThrMet 1292
|||||
4035 ATGTGACGTGTTCCCCAGCATGATGGGCTTTTGGAGAGGCCACCATG 4084
1293 ProGlnAlaProProGlnGlnIlePheProTyrGlnProAsnTyrGlyMetG 1309
|||||
4085 CCACAGCTCTCCGCAACAGATTTCATATCAACCAATTTATGGAATGGG 4134
1309 yGlnGlnIleProAspProAlaPheGlyArgValSerSerProProAsnAlaM 1326
|||||
4135 ACAACAACCAAGATCCAGGCTTTGGTGCAGTGTATGATGCTCCCAATGCA 4184
1326 etMetSerSerArgMetGlyProSerGlnAsnProMetMetGlnHisPro 1342
|||||
4185 TGATGTGTCATGAAGATGGTCCCTCCAGAAATCCATGATGCCAACACCCG 4234
1343 GlnAlaIaIaSerIleTyrGlnSerSerGluMetLysGlyTyrProSerG 1359
|||||
4235 CAGGCTGCATCATCTATACAGTCTCCAGAAATGAAGGGCTGGCCATCAGG 4284

```


[illegible]

714	rserCysGlyAspGlyAsnValValLysGlnGluGlnLeu.....	727
::::::	
12775	TACAGTGCAGCTCGAATTCCTCTTGCGGCGACAGTGAATGGAGCTGCG	12726
728Ser 728	
12725	AATCAAAAGACAAATGGCCTTACCCCTCAGAGAGATCTACTGCATCA	12676
729	ProLysLysGlnuAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAs	745
:::	
12675	CCGCTCAGACAGAAAGACAAATTTAGTCCCTTTCAGATCAGAGATAGGCC	12626
745	pasPProSer.....AspAlaLeuSerLysGluGlnProG 758	
::::::	
12625	TGAGCTTCTCAGTGTATTCAAAGACACACTGTGAACCCCGCCAGAAAGAAA	12576
758	InvAlGluLysAlaAspAsnLysMetSerGlnCysThrSerSerThrIle	774
::::::	
12575	GAGGTGCTCTGGGTATCTCCAGAAACAAAGACCAAAATGTGCTCATTTG	12526
775	ProSerSerSerGlnGlnLysAspProLysIleLysThrGlnThrSerGln	791
:::	
12525	CCACTCTCAAGCCAGATACAGATTAATGAGAGGTGTAGAACATCTCA	12476
791	uGluGlySerGlyAsp..LeuAspAsnLeuAspAlaIleuGlyAspL 807	
:::	
12475	AGAACCCGCGACCAAAATCTGTCTCATTTGTCTTCAGACATTAAGAAA	12426
807	euthSerSerSerSpheryrAsnAsnSerIleSerSerSerSerLysThrHis	823
.....:::	
12425	TGTCACCAAGTAACTT.....GAATCAATCTCTGTGA	12394
824	LeuGlyThrLysGlnGlnValPheGlnGlnThrAsnSerLeuGlyLeuLys	840
::::::	
12393	GTAGAAGAAAGCCCTGCTG.....TCTTTACCTTTGA	12359
840	sSerSerGlnSerValGlnSerIleArgProProTyrAsnArgAlaValS	857
:::	
12358	TCAGAACCCAGTCAACAGGCTTCTTTG.....GAAGCAGTTAG	12324
857	erLeuAspSerProValSerValGlySer.....SerProPro 869	
.....:::	
12323	AGGTCCCTTCAATGGCTCATCTTGGGGGGCCACATTTTCTCCAGAA	12274
870	ValLysAsnIleSerAlaPheProMetLeuProLys.....GlnP 883	
:::	
12273	CATTAAGAAGCTGTCTAATCTCCCACTCAGAGGAGAACAGCTTGGATCAC	12224
883	OMetLeuGlyGlyAsnProArgMetMetAspSerGlnGlnuAsnTyrGlyS	900
:::	
12223	TTTAGAATTTAGAAACTCAGGCCCACTTGCTACAGAAATGATTAAGTGGAT	12174
900	erSer.....MetGlyGlyProAsnArgAsnValThr 910	
:::	
12173	TTTCTCTGAGGTTAAAGAAGATTGGAATGAGCCGTTCTTAATCAGCTG	12124
911	ValThr.....GlnThrProSerSerGln 918	
:::	
12123	GAAGCAGATTCATCTCTAGACATGAAGAACAATAGACAAATCTCTGG	12074
918	YasPTrpGlyLeuProAsnSerLysAlaGlyArgMetGluProMetAsnS	935
:::	
12073	A.....CAAGCAGTTCTGAGTTATCCCA..... 12049	
935	erAsnSerMetGlyArgProGlyGlyAspTyrAsnThrSerLeuProArg	951
:::	
12048	..GATGCACTGTGAAGAAAGGAGGATGTCTTCAATCAGAGCATCTTCA	12001
952	ProAlaLeuGlyGlySerIleProThrLeuProLeuArgSerAsnSerIle	968
:::	
12000	CGTGCTT.....GATGCTGTACCGAAGACACCTCTGAGGAAGAAAGTAG	11954

```

968 eProGlyAlaArpProValleuGlnGlnGlnGlnMetLeuGlnMet 985
    : : : : :
11953 TTCGCACTCTCTCT ..... 11938
985 rProGlyIleIleProMetGlyAlaAsnProTyrglyGlnAla 1001
    : : : : :
11937 .....GAAATGAAGATGTTTACCACAGACCTCA ..... 11908
1002 AlaAlaSerAsnGlnLeuGlySerTrpPro ..... AspglyMetLe 1015
    : : : : :
11907 TCAAGGAGAACAGGTCTGCTCTCCAGGACTTGAAGATGAGG ..... 11863
1015 uSerMetGluGlnIleValSerHisGlyThrGlnAsnArgProLeuLeuArg 1032
    : : : : :
11862 .....TCTGGGACTCTCCCTCGAGGACAGACCTGTCTG 11832
1032 snSer .....LeuAspAspLeuValGlyProProSerAsnLeuGlu 1045
    : : : : :
11831 GGTCTCTCTCGAATGAATATATACCTAGACGCCATCTAGAGGAGGA 11782
1046 GlyIleSerAspGluArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSe 1062
    : : : : :
11781 AGCGAATGTGTAT .....TC 11768
1062 rAsnThrAspAlaThrGlyLeuGluGluIle .....AspArg 1075
    : : : : :
11767 TTCGCCAGAACCGAAAGCTTTGCTCTCAGACTCTCAGGCCGAGAGTCTT 11718
1075 IaleuGlyIleProGluLeuValAsnGlnGlnGlnAlaLeuGluProlys 1091
    : : : : :
11717 CTCGATCATCCCGACAGAGCTCAACAAC .....AGTGTCTTACCCTCCAG 11674
1092 GlnAspAlaPheGlnGlyGlnGlnAlaAlaValMetMetAspGlnLysAl 1108
    : : : : :
11673 AGAGAAAGA .....ACGGGTCTCAGAAATCATCAGTT .....GATCGAAAAAC 11633
1108 a .....GlyLeuTyrglyGlnThrTyrglyProAlaGln 1119
    : : : : :
11632 TGTGGCTCGGACCTCCCTGGGGCAGAGAGTCTCGGATCTCTCCAG 11583
1119 TyProProMetGlnGlyIlePheHisLeuGlnGlnGlnSerProSerPhe 1135
    : : : : :
11582 AACTTGATGTGAACCCAGTGCATCCCTCAGGAAGAGATGATCAAC 11533
1136 AsnSer .....MetMetAsnGlnMetAs 1143
    : : : : :
11532 TCTTCTCCAGATTCTAAAGCCAGACAGAACCCACTTCGCGAGAGAG 11483
1143 nGlnGlnGlyAsnPheProLeuGlnGlyMetHisProArgAlaAsnIle 1160
    : : : : :
11482 TCGGCTCTGGATCATCTCCAGAGTTGACAGCAAACTCTGACTATCCCTC 11433
1160 eTarProArgTyThrAsnThrProLysGlnLeuArgMetGlnLeuGlnGln 1176
    : : : : :
11432 GCGCAGTAGGTCTGGTCTCTCCCTGAAAGTGAAGATAACCAAGACCA 11383
1177 ArgLeuGlnGlyGln ..... 1182
    : : : : :
11382 GCACCCAGGACAGAGTGTCTGATTCTCTCTGAACTAAAGCTTC 11333
1183 .....PheLeuAsnGlnSerArgGlnAlaLeuGlnLeuLeuysM 1135
    : : : : :
11332 AGCCCCCTCGGCCCTTCCAGACAGACGATCAGGTTTCATCAAGCAAG 11283
1195 eGlnAsnProThrIleGlyAlaAlaValMetArgPrometMetGly 1211
    : : : : :
11282 GCAAGAGGCCCTTCTCTGGAAGAGACAGCACTACCGATCTCTCTCTCAA 11233
1211 nProGlnGlnGlyPheLeuAsnAlaGlnMetValAla ..... 1223
    : : : : :
11232 CATCCGCCCA .....AATCCAGAACTCTCTGAGAGGTTCACAGCTC 11192
1224 .....GlnArgSerArgGluLeuLeuSerHisHisPheArgGlnGlnArg 1228

```

```

11191 ATCACCAGAGCCCAAGACCAAGTCTCTACACCACT ..... 11155
1239 ValAlaMetMetMetGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1255
    : : : : :
11154 .....CGAGTCTGAGCTCTCATCATCTCCGGA 11126
1255 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1272
    : : : : :
11125 GCTAACCAAGGAAGGACAGACTGTCCCTGAGAAAGCGCTCTCTCATCT 11076
1272 erPro .....ProProAsnValThrAlaSerProSer 1282
    : : : : :
11075 CACCAAGAACTCGCTTGAACCTCCCTCAGGACACCGGAGAAATCTCTCA 11026
1283 MetAspGlyLeuLeuAlaGlyProThrMetProGlnAlaProGlnGln 1299
    : : : : :
11025 GTG ..... 11023
1299 nPheProTyrglnProAsnTyrglyMetGlyGlnGlnProAspProAla 1315
    : : : : :
11022 .....TCTTCCCGGAGCCAGCCG 11004
1316 .....PheGlyArgValSerSerProProAsnAla 1325
11003 AAAAATCGAGCTCTTCACGCCGAGCGCTCAGCTCATCTCCACGACT 10954
1326 MetMetSerSerArgMetGlyProSerGlnAsnPro 1337
10953 AAGACAACTCTCAGAGAGGCGCTCTCTCTCGCA 10918

seq.name: /cgn2.6/ptodata/2/pna/PCT_NEW_COMB.seq:PCT-US01-01339-7281
seq.documentat_block:
; Sequence 7281, Application PC/TU0101339
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01339
; CURRENT FILING DATE: 2001-03-17
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7281
; LENGTH: 13555
; TYPE: DNA
; ORGANISM: Homo sapiens
; PCT-US01-01339-7281

alignment_scores:
Quality: 226.50 Length: 1330
Ratio: 0.387 Gaps: 67
Percent Similarity: 44.060 Percent Identity: 21.053

alignment_block:
US-09-041-994-2 x PCT-US01-01339-7281 ..
Align seg 1/1 to: PCT-US01-01339-7281 from: 1 to: 13555

330 ThrProValTyArgPheSerLeuAlaAspGlyThrIleValThrAlaGly 346
    : : : : :
5108 TCCCTCAAGTCAAGAAAGATACCAACATTCCTCCAGTCTGATTCTCCACA 5157
346 nThrLysSerLysLeu...PheArgAsnProValThrAsnAspArgHisG 362
    : : : : :
5158 AAACCCAGAGACCTCAGCTGAAAGGGCTTTAGTGTACACCTCTCTGCT 5207
362 TyPheValSerThrHisPheLeu.....GlnArgGlnGlnAsnGly 375
    : : : : :
5208 TATCTCTTCAACATCAGTCTCTCTCTGAGTACCTGTCTCAAAAGACT 5257
376 TyArgProAsnProAsnPro.....ValGlyGlnGlyIle 387

```



```

894  rGlnGluAsnTyrGlySerSerMetGlyGlyProAsn..... 906
      ::::::::::::::::::::
6916  CTCTAAAGACTCACCAGCCAGCAGCATTCCTCCACTCCTCCATCCCA 6965
907  .....ArgAsnVal 909
      ::::::::::
6966  AAGGGGCCCCACTCCCTCAGCTGTGACTCCTCTCTCTCCCAAGAGTA 7015
910  ThrVal.....ThrGlnThrProSerSerGlySptRgIyleuProAs 924
      ::::::::::::::::::::
7016  ACACCTACCCCCCAAGAGACCCCACTCTTCAGTGTGAATCTGCCCTT 7065
924  nSerIleAlaGlyArgMetGlyPro..... 932
      ::::::::::
7066  CCCCAGAGAGGCTCAGCTACTCTCAGCACCACCAAGAGCTCTCTAT 7115
933  .MetAsnSerAsnSer.MetGlyArgProGlyIleAspTyr..... 945
      ::::::::::::::::::::
7116  CCATGACTTCTCTCCCTCCCAAGAGGCGCGAGCACTCCAGCCCTATA 7165
946  .AsnThrSerIleuProArgProAlaIleuGlyGlySerIleProthIleu. 961
      ::::::::::::::::::::
7166  GGAATCCCAAGCTTCCCATCCCAAGAGGCGCCCAACACCCCAAGCTGC 7215
962  .....ProIleuArgSerAsnSerIleProGlyAlaArg..... 972
      ::::::::::::::::::::
7216  AACTCTCCTCCCTCCCTAAAGAGGCGCCCACTACCCCATCCCGAAGTGG 7265
973  .....ProValIleu.....GlnGlnGlnGlnGln 981
      ::::::::::::::::::::
7266  CCCCACACACCCCAAGCTGCACTCCTCCCTCCCAAGAGAGTGCACCT 7315
981  t.....LeuGlnMetArgProG 987
      ::::::::::::::::::::
7316  ACTCCATCCCAAGAGGCGCCCAACACCTCAGCTGCACTCCTCCTC 7365
987  LysIleIleProMetGlyMetGlyAlaAsnProTyrGlyGlnAlaIala 1003
      ::::::::::::::::::::
7366  CCCCAGAGAGGCTCAGCTACTCCATCCCAAGAGGCGCCCAACACCC 7415
1004  SerAsnGlnIleuGlySerTyrProAspIlyMetLeuSerMetGlyGlnVa 1020
      ::::::::::::::::::::
7416  CAGCTGCAACTCTCCTCCCTCCCAAG.....GAACTGCAAGCAGCT 7456
1020  LserHisGlyThrGlnAsnArgProIleuLeuArgAsnSerIleuAspI 1037
      ::::::::::::::::::::
7457  ACCCATTCCTCCCAAGAGGCGCATCA.....CACCCCAAGCTGCAC 7497
1037  euValGlyProProSerAsnLeuGlu.....GlyGlnSerAspIleuArg 1051
      ::::::::::::::::::::
7498  TCCCTCCCTCCCAAGAGGAGTCCAGCAGTACCCCACTCCCAAGAGGG 7547
1052  AlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrG 1068
      ::::::::::::::::::::
7548  CCCCACCAAGAGGCGCATCA...CTCTCCCTCCCAAGAGAGGTCA 7594
1068  LLeuGlnGlnIleAspArgAlaIleuGlyIleProGlyIleuValAsnG 1085
      ::::::::::::::::::::
7595  GCTACCCCATCTCAAGAGGCGCCCACTCCCAAGCTG..... 7634
1085  LysGlnAlaLeuGlnProLysGlnAspAlaPheGlnGlyGlnIleAla 1101
      ::::::::::::::::::::
7635  .CGACCTCTCTCCCTCCCAAGAGGCG..... 7661
1102  ValMetMetAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAlaG 1118
      ::::::::::::::::::::
7662  .....CAGCTACCCCAATCCCAAG 7681
1118  nGlyProPro.Met..... 1122
7682  GGGGCCCCCATGCCCCCAGCTGCAACTCCTCCCTCCCAAGAGAGTCT 7731

```

```

1123  .....GlnGlyGlyPheHisLeuGlnGlnGlnSerProse 1134
      ::::::::::::::::::::
7732  AGCTACCCCAACCCCAAGAGGCGACCCAC.....AACCCAG 7769
1134  rPheAsnSerMetLeuAsnGlnMetLeuGlnGlnGlnLysAsnPheProLeuG 1151
      ::::::::::::::::::::
7770  CTGCACACTCTCTCCCTCCCAAGAGAGTCTAGTACCCCAACCCCAAG 7819
1151  LnglyMetHis...ProArgAlaAsnIleMetArgProArgThrAsn... 1165
      ::::::::::::::::::::
7820  GGGGCCCCCAACCCCAAGTGCACACTCTCTCCCTCCCAAGAGAGTCT 7869
1166  .....ThrProLys..... 1168
      ::::::::::::::::::::
7870  AGCTACCCCAACCCCAAGAGGCGCCCAACCAACCCCAAGTGCACCTC 7919
1169  .....GlnLeuArgMetGlnLeuGlnGlnArgLeuGlnGly 1181
      ::::::::::::::::::::
7920  CTTCACCAAGAGAGTCTAGTACCTACCCCAAGAGGCGCCCAACCA 7969
1181  LnglnPheLeuAsnGlnSer.....ArgGlnAlaLeuGlnLeu 1193
      ::::::::::::::::::::
7970  ACCCAAGCTGCACACTCTCTCCCTCCCAAGAGAGTGTGCTACCCATC 8019
1194  LysMet.....GluAsnProThrAlaGlyGlyAlaAlaValMetArg 1207
      ::::::::::::::::::::
8020  CCCCAGAGGCGCCCAACCAACCCCAAGTGCACACTCTCTCCCTCCCA 8069
1207  g.....PrometLeuGlnProGlnGlnGlnGlyPheLeu 1218
      ::::::::::::::::::::
8070  GAGGCGCCAGTACCCCAACCCCAAGAGGCGCCCACTCCCAAGCTCA 8119
1218  snAlaGlnMetValAlaGlnArgSerArgGlyLeuLeuSerHisHisPhe 1234
      ::::::::::::::::::::
8120  ACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8169
1235  ArgGlnGlnArgValAlaMetMetGlnGlnGlnGlnGlnGlnGlnG 1251
      ::::::::::::::::::::
8170  CCCCATCTCCCAAGTGTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8219
1251  nGlnGlnGlnGlnGln.....GlnGlnGlnGlnGlnGlnGlnG 1264
      ::::::::::::::::::::
8220  ACTTCACCCCAAGAGGCGCCCACTCTCTCTCTCTCTCTCTCTCTCT 8269
1264  LnglnGlnGlnThrGlnAlaPheSerProProProAsnValThrAlaSer 1280
      ::::::::::::::::::::
8270  CCCCAGAGAGGCGCGAGGTACCCCAACCCCA..... 8301
1281  ProSerMetAspGlyLeuLeuAlaGlyProThrMetProGln..... 1294
      ::::::::::::::::::::
8302  .....AAGGCGCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8330
1295  AlaProProGln.GlnPheProTyrGlnProAsnTyrGlyMetGlyGln 1310
      ::::::::::::::::::::
8331  CCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8380
1311  GlnProAspProAlaPheGlyArgValSerSerProProAsnAlaMetMe 1327
      ::::::::::::::::::::
8381  TAAAGAGAGGCGCACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8425
1327  tSerSerArgMetGlyProSerGlnAsnPrometMetGlnHisProGln 1344
      ::::::::::::::::::::
8426  .....CCAGCTGCACACCCCTC.....CCACAA 8450
1344  IalaSerIleTyrGlnSerSer..... 1351
      ::::::::::::::::::::
8451  GAGAGTCCCTATGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8500
1352  .....GluMe 1353
8501  ATCTCCCAAGAGGAGCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 8550
1353  tLysGlyTyrProSerGlyAsnLeuAlaArgAsnSerSerPheSerGln 1370

```

```

8551 AAAAGCTCCAGCACTCCAGTCACGAGAAAGCGCAGCCACCCATCC 8600
1370 InGlnPheAaHISGInGlyAsnProAlaValTySerMetValHISmet 1386
8601 AAAGAGATCTCAC.....TCCCCAGCAGTGCCTGTCCTCCCTCAA 8644
1387 AsnGlySerSerGlyHISmetGlyInMetAsnMetAsn..PrometPro 1402
8645 AAAGGCCCCCAACTTCAGCCCCCAAGAGGCCAGCCACCCCATCTCT 8694
1403 MetSerGlyMetProMetGlyProAspGlnIlys 1413
8695 CCAAGGGAGATCCCA.....CCCTCCAGCAG 8721

seq_name: /cgn2_6/ptodata/2/pna/US09_NFW_COMB.seq:US-09-335-032-12207

seq_documentation_block:
; Sequence 12207, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335.032
; PRIOR FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12207
; LENGTH: 1531974
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; US-09-335-032-12207

alignment_scores:
Quality: 226.00 Length: 1299
Ratio: 0.400 Gaps: 61
Percent Similarity: 43.495 Percent Identity: 19.477

alignment_block:
US-09-041-994-2 x US-09-335-032-12207/rev ..
Align seg 1/1 to reverse of: US-09-335-032-12207 from: 1 to: 1531974

190 GlnTySerHISThrPheAsnGly.....ArgMetLeuMetIly 202
::: ::::: ::::: ::::: ::::: :::::
1195132GAGCGCGGACATACAGTATTTGCGCGCGTAACTACAAATTTGTGTTCT 1195083
202 sThrProHISAspIleLeuGlnAspIleAsnAlaSerProGlnMetArg 219
::: ::::: ::::: ::::: :::::
1195082TCCCTTCGAAACCTTAGAAGGAGTA..... 1195055
219 InArgTyGlnThrMetGlnGlyPheAla..LeuSerGlnProArgAla 234
1195054.....AGATTTTTCGACTTGTGCGAGCGCGCAAAAT 1195022
235 MetMetGlnGlnGlyLysAspLeuGlnSerCysMetIleCysValAlaIar 251
::: ::::: ::::: ::::: :::::
1195021TTTCAAGG.....CGGCTATTACCCGTAGCTTAGCCG 1194987
251 GArgIle.....T 254
1194986CCGAGACAGATAGATATAAGATTATTAGACATCATATCAAAAGTA 1194937
254 hTrHcGlyGlnArgThrPheProSerAsnProGlnSerPheIleThrArg 270
::: ::::: ::::: ::::: :::::

```

```

1194936CACTTCCCGCTAAAGATGCTCTCACTCCAGATTCAGACGTGCT... 1194890
271 HISAspLeuSerGlyLysValAlaAsnIleAspThrAsnSerLeuArgse 287
1194889.....AATTCACATCATCTCAGTGAAGACCAAGCATGTATGAG 1194849
287 rSerMetArgProGlyPheGlnAspIleIleArgArg..... 299
1194848TAAATTT...CCTAAGCAGATGACTTAATTGATGAAGAGATGAAGAA 1194802
300 .....CysIleGlnArgPhePheSerLeuAsnAspGlyGln 311
1194801TGACCCCACTATATCTGCTATCCGCTTAACCCAGCTGCTAGATTAA 1194752
312 SerTrpSerGlnLysArgHISArgInGlnAlaTyLeuAsnGlyHIS.. 327
1194751GATGAAGAAACAACTTCGGAAGAAATAGATGCATTTTAAAGAAACGA 1194702
327 ..... 327
1194701TATACGAAGGGGTATAGGTTTGACGAGCGTCACAGCCGAACCTCTAC 1194652
328 ....AlaGlnThrProValTyArgPheSerLeuAlaAspGlyThrIle 342
1194651ATACACCTGCGACTCCCATPAACAAAGAAATGAGATGTAATTTTA 1194602
343 ValThrAlaGlnThrLys.....SerTyLeuPheArgAsnPr 355
1194601ATTAATGTACCTACCAAGCGCTATCCGATTTCAAAATGATGGAGACT 1194552
355 oValThrAsnAspArgHISGlyPheVal.....SerThrHISpHeLeuG 370
1194551GCCGAAATATAGCGAAATGTATCACTCAAAATCAGAGACTTTTGTTC 1194502
370 InArgGlnGlnAsnGlyTyArgProAsnProAsnProValGlyGlnGly 386
1194501CAAGCCACGATATTCATTCCTGTAAT.....ATGATGATTTCT 1194461
387 IleArgPro.....PrometAl 392
1194460TTGAGCGCGCTGAAGACAGAGGATATGTGACCAAAAGCGCACTT... 1194413
392 aGlyCysAsnSerSerValGlyGlyMetSerMetSerProAsnGlnGly 409
1194412.....GAGTCACATATGAGGAGTGAAGAGAGCAGACATGAAGAAAGAC 1194370
409 euGlnMetProSerSerArgAlaTyGlyLeuAlaAspProSerThrThr 425
1194369AGAGATG..... 1194362
426 GlnGlnMetSerGlyAlaArgTyGlyGlySerSerAsnIleAlaSerIe 442
1194361...CAATTCACATTCATGGAATCTCTAATTAATTCGAATAGCTTCAT 1194315
442 uThrProGlyProGlyMetGlnSerProSerSerTyGlnAsnAsnAsn 459
1194314ACCTATTTGCGCGCGCTCCCGCACTGCTTAATGATTAAGACACTACACT 1194265
459 TrpGlyLeuAsnMetSerSerProProHISGlySerProGlyLeuAlaPro 475
1194264ATTATATGAGACATGAGTACCCAGACAGATCCCACTAGAGGCAACCT 1194215
476 AsnGlnGlnAsnIleMetIleSerProArgAsnArgGlySerProIlysil 492
1194215..... 1194215
492 eAlaSerHISGlnPheSerProValAlaGlyValHISSerPrometAlas 509
1194214..GCAGTAAATTCAGATCCGTTG.....ATTAAGAACATTTATGGCGCA 1194172
509 eSerSerGlyAsnThrGlyAsnHISSerPheSerSerSerIleuSerAla 525
1194171AAGAGATAGATATCATCAGACAT...TATACGAGAGAAATGTTGACGCA 1194125

```

```

526 LeuGlnAlaIleSerGluGlyValGlyThrSerLeuSerThrLeuSe 542
      |||
      |||
      |||
1194124TTAGAACCGTA.....TTTTATTA..... 1194104
542 rSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleThrGlnP 559
      |||
      |||
      |||
1194103.....ATGACGATTATATAC 1194088
559 rSerLysValSerAsnGlnAspSerLysSerProLeuGlyPheTyrcys 575
      |||
      |||
      |||
1194087CAATCGAAATATACCCAGGCTTTCCTTAGCAAAATGGAATGCATGTC 1194038
576 AspGlnAsnProValGlnSerSerMetCysGlnSerAsnSerArgAspH 592
      |||
      |||
      |||
1194037CAAACTCTGCCT..... 1194026
592 sLeuSerAspLysGlnSerLysGlnSerSerValGlnGlyAlaGlnAsnG 609
1194026..... 1194026
609 LmArgGlyProLeuGlnSerLysGlyHisLysLysLeuGlnLeu 625
      |||
      |||
      |||
1194025.....CTATTAATCTAATA 1194011
626 ThrCysSerSerAspAspArgGlyHisSerSerLeuThrAsnSerProle 642
      |||
      |||
      |||
1194010.....CCGAGAGCTCACAAAGGCTTGAACCACTAATATTAT 1193976
642 uAspSerSerCysLysGlnSerSerValSerValThrSerProSerGly 659
      |||
      |||
      |||
1193975AAACACAGCGCTTAACAGAGCCGAGATTACTGTGTTGGC...TCCGCA 1193929
659 aLsSerSerThrSerGlyGlyValSerSer..... 669
      |||
      |||
      |||
1193928TGGAGAACACTAGAACACTGTGATTGTGCTTAAACACCTAAAGA 1193879
670 .....ThrSerAsnMetHisGlySerLeuLeuGlnG 680
      |||
      |||
      |||
1193878TTCATGATCCGCGAAACACACAAATACCGCATCAAAACATCTTTAGA 1193829
680 uLysHisArgGlyLeuHisLysLeuGlnAsnGlyAsnSerProAlaG 697
      |||
      |||
      |||
1193828AGAAAGCTAAATGATGCAAGCAGATTTTAAAGAGGCCACAAA...TATA 1193782
697 LuValAlaLysIleThrAlaGlnAlaThrGly...LysAspThrSerSer 712
      |||
      |||
      |||
1193781AACTACTATCTGCACAGCAATGCGACAGGCAATCAAGGAC..... 1193741
713 IleThrSerCysGlyAspGlyAsnValValLysGlnGlnGlnLeuSerP 729
      |||
      |||
      |||
1193740TACTGACCTATGCGGAAATATGCTGTGTAACGTAACACTTTAGC 1193691
729 oLysLysLysGlnAsnAlaLeuLeuArgTyLeuLeuAspArgAspA 746
      |||
      |||
      |||
1193690TGGTAAAGAA.....AATTAATTA.....TCGACG 1193665
746 SP.....ProSerAspAlaLeuSerLys 753
      |||
      |||
      |||
1193664ACCGCGCATATCTGAAGTCCGTAAGACCATCGATCTTCAGAAAT 1193615
754 GlnLeuGlnProGlnValGlnGlyValAspAsn.....Lys 765
      |||
      |||
      |||
1193614GATTCGATTAATGACATAGCCGTAAGGATGACATTCGAAATTAATTCGAA 1193565
765 sMetSerGlnCysThrSerSerThrIleProSerSerSerGlnGlnLysA 782
      |||
      |||
      |||
1193564CGTTGACGATTAATCT.....GAAAGAG 1193542
782 sProLysLysIleLysThrGlnThrSerGlnGlnGly..... 793
      |||
      |||
      |||
1193541AAAGTCTGCTGCTAAGCATTAATGATGAGAAATGAAAGATGAAGCAGCA 1193492

```

```

794 ....SerGlyAspLeuAspAsnLeuAspAlaIleLeuGlyAspLeuThrSe 809
      |||
      |||
      |||
1193491GCTTAAGAGCATTTGATTTTGCAGATGGTTTGTATCTCAAGAGGTGC 1193442
809 rSerAspPheTyAsnAsnSerIleSer..... 818
      |||
      |||
      |||
1193441CCAGCATCAAAATATATCTCTATGTCACACCAAGTTACTCTAATAAAG 1193392
819 ..SerAsnGlySerHisLeuGlyThrLysGlnGlnValPheGlnGlyThr 834
      |||
      |||
      |||
1193391CATCTCCCTCCCAACGGAAGTGTGTTTAATCAACATCAAGTAGCTGATCA 1193342
835 AsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleArgProP 851
      |||
      |||
      |||
1193341TCAGCGTTG...ATCGAAACCGAAGATCAAGAAAGAGCTTACCCCCC 1193295
851 oTyAsnArgAlaValSerLeuAspSerProValSerValGlySerSerP 868
      |||
      |||
      |||
1193294ATTAAACTTTCAACTTGTTCACGAA..... 1193267
868 roProValLysAsnIleSerAlaPheProMetLeuProLysGlnProMet 884
      |||
      |||
      |||
1193266.....CTCAATCTTTGAAAAAACAATTAATACAGGATTACCGCTG 1193225
885 LeuGlyGly.....AsnProArgMetMetAspSerGlnGlnAs 897
      |||
      |||
      |||
1193224TTATATAGTATTTATGACAGCGCCCAAAAAAGATGATTTCTTACCANT 1193175
897 nTyGlySerSerMetGlyGlyProAsnArgAsnValThrValThrGln 914
      |||
      |||
      |||
1193174TATTCCAATTTCCAA.....TCGGTGTGTCAC 1193146
914 hrProSerSerGlyAspTrpGlyLeuProAsnSerLysAlaGlyArgMet 930
      |||
      |||
      |||
1193145TGTGATATATGATTTTATTAAGTTTACTTGACGCTCAATGATTCAGCA 1193096
931 GlnProMetAsnSerAsnSerMetGlyArgProGlyLysArgTyAsnTr 947
      |||
      |||
      |||
1193095GAACCAATCAATATACAAATTCAGCAAAAGCGGGATGTGTTAT..... 1193051
947 rSerLeuProArgProAlaLeuGlyLysSerIleProThrLeuProLeuA 964
      |||
      |||
      |||
1193050.....GGTATAGACAAATCATTAATTAA 1193026
964 rSerAsnSerIleProGlyAlaArgProValLeuGlnGlnGlnGln 980
      |||
      |||
      |||
1193026GACACACAGCATTCCTCATTTAGCG...TATCTCAAAATATACAAACCG 1192979
981 MetLeuGlnMetArgProGlyGlnIleProMetGlyMetGlyValAsnTr 997
      |||
      |||
      |||
1192978ACTATATGCTGTGTGAGACGATCAAGAACTGTCAAAAATATTAAAC 1192929
997 oTyGlyGlnAlaAlaIleSerAsnGlnLeuGlySerTrpProAspGly 1014
      |||
      |||
      |||
1192928ATATGCTAC.....AATGCG..... 1192913
1014 etLeuSerMetGlnGlnValSerHisGlyThrGlnAsnArgProLeuLeu 1030
      |||
      |||
      |||
1192912.....GAATGATTAAGTGCACATATGACATATGSC...CTGCAC 1192877
1031 ArgAsnSerLeuAspAspLeuValGlyProProSerSerSerLeuGly.... 1045
      |||
      |||
      |||
1192876TATTTCTAATTA.....TCGAAATTAATCAAGGTAG 1192848
1046 .....GlyLysSerAspGlnArgAlaLeuLeuAspArgGlnIleThr 1060
      |||
      |||
      |||
1192847AAGCCATGCGCAAGTGTGAGCGTTTGTGA.....CAATTAAGAGAAC 1192804
1060 euLeuSerAsnThrAspAlaThrGlyLeuGlnGlnIleAspArgAlaLeu 1076
      |||
      |||
      |||
1192803GATTCATATTTTACTGACSTCAAGGT.....CCCAAGCAAT 1192766
1077 GlyIleProGlnLeuValAsnGlnGlnAlaLeuGlnIleProLysGlnAs 1093

```

```

1192761GCTGCTCAACAGTGGTGTGATTGAAGCCCATTAATTTCAGCAGAGACAG... 1192718
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1192762GCTGCTCAACAGTGGTGTGATTGAAGCCCATTAATTTCAGCAGAGACAG... 1192718
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1093 pAlaPheInglngIngluAlaIaIaValMetMetAspGlnLysAlaLyl 1110
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1192717..... 1192702
1110 eutyrgIyGlnthrtTyrrProIaInglngIyProPromeInglngIyGlnPhe 1126
      ||| :::: :::: :::: :::: :::: :::: :::: :::: ::::
1192701CTTTAGCT.....GTAATACAGAAATCCATACACACAGGTCAT 1192664
1127 HsLeuInglngIyGlnSerProSerPheAsnSerMetMetAsnInglMetas 1143
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1192663AGAGACACTCCGTTGGCGCAGCATGTGTCAGCGCATTTGAAAAGATGATGAA 1192614
1143 nGlnInglngIyAsnPheProLeuInglngIyMetHisProArgAlaAsnIleu 1160
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1192613AAAAAGAGAAAT.....ACGCGCGATCCCAATCCAA 1192582
1160 eArgProArg.....ThrAsnThrPro 1167
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1192581CCCAAGCAAGAAACCGTTAGATTGTAAAAACATGAAGATCCCTACCCCA 1192532
1168 Lys..... 1168
1192531GCAGAAATGCTCGTTTGAAGACCCAAAGAGATGAGCATTTGAGAAAGCA 1192482
1169 ... GlnLeuArgMetGlnLeuGlnInglngIyGlnInglngIyGlnPheI 1184
      ||||| :::: :::: :::: :::: :::: :::: :::: :::: ::::
1192481TTTACCAATTGAGAAAGACAGTGAATAAGATTTCACACAGAGGCA... 1192436
1184 euaSngInSerArgGlnIaIaLeuGlnLeuLysMetGluAsnProThrAla 1200
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1192435...CACGAAGTACGACAGGCACATTTCGCCAGTCAAAAGCCAAATCCCT 1192388
1201 .....GlyAlaIaAlaIa 1205
1192387TCMAATGGCAATCTCTTCAAAATTTGGCTGAATGGCAAGAGCTCGGC 1192338
1205 lMetArgProMetMetGlnProGlnInglngIyPheLeuAsnAlaGlnMetV 1222
      ||||| :::: :::: :::: :::: :::: :::: :::: :::: ::::
1192337TCCAAAGCA...AACCGAAACAGTAACTATGACGACAGATATCA 1192297
1222 aIaIaGlnArgSerArgGlnLeuLeuSerHisHis..... 1233
1192296TTGAAGAGTTACTCGAGAAAGTTGTTGGAGCAAAAACAGACATGTCCT 1192247
1234 .....PheArgGlnI 1237
1192246GAGATGCCCTTAAGCGCGCAAAAACATATTATGAACCTTGAAGAGCA 1192197
1237 nArgValAlaMetMetMetGlnInglngInglngInglngInglngInglng 1254
      ||:: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1192196ACGACACAGCTTAAGACGACGCAATTCACACACAAAGGAGCAGTTAC 1192147
1254 InglngInglngInglngInglngInglngInglngInglngInglngInglngAla 1270
      ||:: :::: :::: :::: :::: :::: :::: :::: :::: ::::
1192146AGAGAGAGTCGATCGACGACGCAATTCGACGAATACAAACACAGGCTCT 1192097
1271 PheSer.ProProProAsnValThrAlaSerProSerMetAspLysLeu. 1286
      ||| ||||| :::: :::: :::: :::: :::: :::: :::: ::::
1192096TCAAGCTCGCCACGAA...TCATTCATCTCAATCAATCTTCTTT 1192053
1287 .....LeuAlaGlnProThrMet 1292
1192052CCAAATTTCTTACCTTAATTCTGCTCCAGAGATAAAGTCAACCACT... 1192006
1293 ProGlnAlaProProGlnInglngPheProIyGln...ProAsnTyrr 1306
      ||||| ||:: :::: :::: :::: :::: :::: :::: :::: ::::
1192005CCACAGGAGATTTTCCAAAGGTTCCAAAACAAATAGACTTAACCTT 1191961
seq_name: /cgn2_6/protdata2/pna/US09_NEW_COMB_seq:US-09-335-032-12216

```

```

seq_documentation_block:
- Sequence 12216, Application US/09335032
- GENERAL INFORMATION:
- APPLICANT: Velculescu, Victor
- APPLICANT: Vogelstein, Bert
- APPLICANT: Kinzler, Kenneth
- TITLE OF INVENTION: Characterization of the yeast
- TITLE OF INVENTION: Transaptome
- FILE REFERENCE: 01107.78572
- CURRENT APPLICATION NUMBER: US/09/335,032
- CURRENT FILING DATE: 1999-06-16
- PRIOR APPLICATION NUMBER: US 60/035,917
- PRIOR FILING DATE: 1997-01-23
- PRIOR APPLICATION NUMBER: US 09/012,031
- PRIOR FILING DATE: 1998-01-22
- NUMBER OF SEQ. ID NOS: 12219
- SOFTWARE: FastSeq for Windows Version 4.0
- SEQ ID NO 12216
- LENGTH: 924430
- TYPE: DNA
- ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12216

```

```
alignment_scores:
  Quality: 224.00      Length: 1626
  Ratio: 0.319         Gaps: 75
Percent Similarity: 43.235      Percent Identity: 19.619
```

alignment_block:

US-09-041-994-2 x US-09-335-032-12216/rev .

Align seg 1/1 to reverse of: US-09-335-032-12216 from: 1 to: 924430

24 ThrProGlyGlnGlyLeuThrCysSerGlyGluLysArgArgGlu 400

258 ACCGAGGTGAAGT.....GAAGTGAAATGAAATAAGAGAA 59

40 nGluSerLysTyrIleGluGluLeuAlaGluLeuIleSerAlaAsnLeuS 57

217 GGAAGAAGTCTAGAGGAA.....C 59

57 eraspileaspaspnpheasnvallysprowsplyscysalalleulys 73

194 CAGATTTT TAGATATCAATGACACTGGAGATTCAGGCGAGTAAAAATAGT 59

74 GluThrValArgGlnIleArgGlnIleLysGluGlnGlnGlyLysThrIleLeu 90

144 GAAACTACAGGTCGCTTCT...CCTACTGATCCTCTACACGAC 59

90 rAsnAspAspValGlnLysAlaAspValSerSerThrGlyGlnGlyV 100

100 ACCACCAGATTCAGTTGAGAGAGG..... 59

107 a)IleAspLysAspSerLeuGlyProLeuLeuLeuGlnAlaLeuAspGly 122

075AAGCTTGTGAGCAGCGTACAAAGAAC 59

1124 PheLeuPheValValAlaSerGluAlaAsnIleValPheValSerGluAs 14

048 TTTCAGCAAAACTTAACAACAACTCACTGAATC.....CA 59

140 nValThrGlnTyrLeuGlnTyrLysGlnGluAspLeuValAsnThrSer 15

10 TATACCAGTAACTTGTATACATGCGGACTTACTACTAGACACTGAAG 59

157 aTyr.....AsnIleLeuHISGluGluAspArgLysAspPhe 16

560 GTTACAAGCCAAACAGATATGTCATGACACAGACATACAA 59

170 [ell]vSAsnI[ell]ProIvS[er]ThvYa)A[ng]IvYa)S[er]Tr[ro]ThA[ng]I 178

016 CTGAGGCAGTTCATTAAGAACCACAAATCAGCAATAATCATC **58**

590444 AATGTCC.....AAAGGGACATTTGAAAAACAATATATT 590410
727 eu.....SerProLysLysSLuSnAsnAla 736
590409 TCGAGCTAGTTCCTCCACTTCGATCGCCAAA..... 590378
737 LeuLeuArgTyrLeuLeuAspArgAspProSerAspAlaLeuSerly 753
590377 ..CTAAGATATTATTAGACGATATGATGATGAAGACTCCAAATTAC..... 590336
753 sGLuLeuGlnProGlnValGluGlyValAspAsnLysMetSerGlnCysT 770
590335AATGAACCTAAATCATAGCA 590317
770 hrSerSerThrIleProSerSerSerGlnLulLysAspProLysIleLys 786
590316 GAGCATCTCTCTCAGCGCGGACTAGATATGATTAAGAGTTCAA...AAG 590270
787 ThrGluThrSerGluGluGlySerGlyAspLeuAspAsnLeuAspAlaI 803
590269 CTCTCTGTCTCAAGGCCGCGCTCCA.....ACAAGAAAT 590235
803 eLeuGlyAspLeuThrSerSerAspPheTyrAsnAsnSerIleSerSerA 820
590234 TCTTTCAAAGATCAACCTTTCAAATGACGTACATTCTGCGGAATTCAG 590185
820 sn.GlySerHisLeuGlyThrLysGlnGlnValPheGlnGlyThr.... 834
590184 ATACCGAATCACGAAGGAGCAAAAGTTAATGAACAGCGGCATATCAG 590135
834 834
590134 CTACCCACAACTCAATGACGACATAAAGTTTGATTCAAGATTAAGCA 590085
835 ..AsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleArgPro 850
590084 TCAATTTCCGTGACGATGATTCGACGACAGATCAATTCAGAGTTTAAGAT 590035
851 ProTyrAsnArgAlaValSerLeuAspSerProValSerValGlySerSe 867
590034 TCTTGAGCGCGAATATTATT.....ATTAAGATCCATATCTGC 589997
867 rProProValLysAsnIleSerAla.....PheProMetLeuP 880
589996 AATTAAAGTAAATTAATGAAGTACTAAATTAATTTTCAATTATTATT 589947
880 rOLysGlnProMetLeuGlyGlyAsnProArgMetMetAsp..... 893
589946 AAGAAAGAGAAAGAAAAAAAGACAAATTTTTCATGAAAGAAAA 589897
893 893
589896 AGCTAAACGATGATTGTACGTACCAAGTCTCTCTGAGCTATTTCGTA 589847
894SerGlnGluAsnTyrGlySerSerMetGlyGlyProAsnA 907
589846 ATTACATACCTTCTCGAAATGTCAAAGTAATATCTTCTATACGACAC 589797
907 rGAsnValThrValThrGlnThrProSerSerGlyAspTrpGlyLeuPro 923
589796 GAAACTCGTTCCTGCAACAGGTATTACGTACGTGA..... 589759
924 AsnSerLysAlaGlyArg.....MetGluProMetLysSerAsnSerSe 938
589758AAAGGGCCCAAGAGGCATATGCCAATATCAATGCTGTATAC 589715
938 tGlyArgProGly.....GlyA 944
589714 AGGAGAGAGAGGCGTATATAATACCATAGTAGAAAAAGACCAAG 589665
944 sPTyrAsnThrSer.....LeuProArg 951

589664 ATTTCTATTGCAATCCTGTGACCTCTTTGGTTCCTTCCCTTGGTATCT 589615
952 ProAlaLeuGlyGlySerIleProThrLeuProLeuArgSerAsnSerI 968
589614 CCC.....GTTCTTTTGTCAATTTTTCGTAACTGACGGAAT 589574
968 eProGly.....AlaArgProValLeuGln 977
589573 AAGAAAAAGGTGCATGATTAAGATATGATATACAGACCAATATTAATAC 589524
977 InGlnGlnGlnMetLeuGlnMetLArgProGlyLulIleProMetGlyMet 993
589523 AATGAGCGCTGTAGTACACGCGATTTTAGCTCAGCTCCCAATATCAAA. 589475
994 GlyAlaAsnProTyrGlyGln..... 1000
589474 GGTTCGAATAGGCATACCTACGTATGCGAGTTTGATGGGATTCGAGTTT 589425
1001AlaAlaAlaSerAsnGlnLeu..... 1007
589424 CGTTTCGATGTATATGCTTAAGCACTTAAGCAATATATTGTACGCCATA 589375
1007 1007
589374 TTATTAATATTATTAACAAAAACAATTACTGCACTGCGCCAAATTT 589325
1008GlySerTrpProAs 1012
589324 TTAAAGCAAGCTGACCTTCCTTATCTAGATAAAGCATGCTGCTCGG 589275
1012 pGlyMetLeuSerMetGluGlnValSerHisGlyThrGlnAsnArgProL 1029
589274 TGGAAAAACTAGCTTG.....AACGCCAGCC 589249
1029 euleuLeuArgAsnSerLeuAsp.....AspLeuValGlyPro 1040
589248 TTAAGCAGGATTAATGATATTCATCTTAAGCGTGACATTTTGTGTA 589199
1041ProSerAsnLeuGlnGlyGlnSerAspGluArgAlaLe 1053
589198 GATGGCGTATTACCTTCGAAATATGCTGATGACGCTAATGACACCTTTT 589149
1053 uLeuAsp.....GlnLeuHisThrLeuLeuSerAsnThrAsp...AlaT 1067
589148 ACTGGAATGCTGGAAATTTTTCATATCAATCTTTCATAGACACCTAGAT 589099
1067 hrGlyLeuGlnGluIleAspArgAlaLeuGly.....IleProGlu 1080
589098 CTGGGTACCAACAGATCTAATCTTTAAGACAGAGATTAATACCAATT 589049
1081 LeuValAsnGlnGlyGlnAlaLeuGlnLupProLysGlnAspAlaPheGlnI 1097
589048 TTGCCAGCTAATTTCTAAATCTAATATGCTTCCTCCAT..... 589013
1097 yGlnGlnAlaAlaValMetMetAspGlnLysAlaGlyLeuTyrGlyGlnT 1114
589012TTCCTCTAATC 589003
1114 hrTyrProAlaGlnGlyProProMetGlnGlyGlyPheHisLeuGlnGly 1130
589002 TCCCAACCAATGTGATTCACCACTCAAAATATGTTTCCAGTTTCAGAG 588953
1131 GlnSer.....ProSerPheAsnSerMetLysGlnMetLysGlnGlnI 1145
588952 GAGAGTTTAGGCCCAATGCGTACAGGAAGTACTTAATTATTAACAGATC 588903
1145 nGlyAsnPheProLeu.....GlnGlyMetH 1154
588902 AACCAATCGAAGAGCTCTCGAAGATTTCTATCGAAGATTCGCGGTCT 588853
1154 IsProArgAlaAsnIleMetLArgProArgThrAsnThrProLysGlnLeu 1170
588852 ACGACAAACAGATAGCCCTAATTTTGCACCTGATACT..... 588815

```

1171 ArgMetGlnLeuGlnGlnArgLeuGlnGlnGlnPheLeuAsnGln 1187
      :::::::::::::::::::: ||||| ::::::::::::::::::::
588614 .....GCTTTAAACAGTGAATTGCTGGGCGAGCATACCACTATPAA 588771
1187 rArgGlnAlaLeuGlnLeuLysMetGluAsnProThrAlaGlyAla 1204
      :::::::::::::::::::: ||||| ::::::::::::::::::::
588770 TCTACATTAACACTTCATGATTTGCATCACCAGCA..... 588734
1204 lAlaMetArgProMetMetGlnProGlnGlnGlyPheLeuAsnAlaGln 1220
      :::::::::::::::::::: ||||| ::::::::::::::::::::
588733 .....CAGCCCGAGCATCA.....TCTCACAG 588710
1221 MetValAlaGlnArgSerArgGluLeuLeuSerHisHisPheArgGln 1237
      :::::::::::::::::::: ||||| ::::::::::::::::::::
588709 CAGTTCAGACGAGCTCGAGCATCA.....CCCAGCATCAACCGCAGCA 588666
1237 rArgValAlaMetMetMetGlnGlnGlnGlnGlnGlnGlnGlnGln 1254
      :::::::::::::::::::: ||||| ::::::::::::::::::::
588665 GCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCAAC 588616
1254 lGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnAla 1270
      :::::::::::::::::::: ||||| ::::::::::::::::::::
588615 AGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCA 588566
1271 PheSerProProAsnValThrAlaSerProSerMetAspGlyLeu 1287
      :::::::::::::::::::: ||||| ::::::::::::::::::::
588565 CAGACACCGATCTCTATTGTCAAC..... 588542
1287 uAlaGlyProThrMetProGlnAlaProProGlnGlnPheProTyrGlnP 1304
      :::::::::::::::::::: ||||| ::::::::::::::::::::
588541 .....CCACAAATGCTCCT..CATATTCATCGGAAA 588511
1304 rAsnTyrGlyMetGlyGlnGlnProAspProAlaPheGlyArgValSer 1320
      :::::::::::::::::::: ||||| ::::::::::::::::::::
588510 ATTCTCATTCACCGGACTATGCT.....TCA 588482
1321 SerProPro.....AsnAlaMetMetSerSerArgMetG 1332
      :::::::::::::::::::: ||||| ::::::::::::::::::::
588481 GTGCTCTCTCAAAATCAACAATTTAAATGCAAAACCAATCTTCAAATGTT 588432
1332 yProSerGlnAsnProMetMetGlnHisPro..... 1342
      :::::::::::::::::::: ||||| ::::::::::::::::::::
588431 TTCAGACCAGAGCGCTTTCATTAATTCATTAATTCACCAACAAATCAAG 588382
1343 ..GlnAlaAlaSerIleTyrGlnSerSerGlyMetGlyTyrProSer 1358
      :::::::::::::::::::: ||||| ::::::::::::::::::::
588381 GACAGGCGCATCTTTTCAGCAAGCCAA.....TCT 588350
1359 Gly.....AsnLeuAlaArgAsnSerSerPheSerGlnGlnP 1372
      :::::::::::::::::::: ||||| ::::::::::::::::::::
588349 GGCAGGTTTGATGACATGAAAGCTATGAAATGTTTTCAGCGCAGCAAGC 588300
1372 eAlaHisGlnGlyAsnProAlaValTyrSerMetValHisMetAsnGly 1389
      :::::::::::::::::::: ||||| ::::::::::::::::::::
588299 ACTACAGCAA.....AATTCAC 588283
1389 eSerGlyHisMetGlyGlnMetAsn 1397
      :::::::::::::::::::: ||||| ::::::::::::::::::::
588282 TACAGCAAAATCTTGAAATCAAAAT 588257
seq_name: /sgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12217
seq_documentation_block:
; Sequence 12217, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velulescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032

```

```

; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12217
; LENGTH: 784328
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12217

alignment_scores:
  Quality: 218.50      Length: 951
  Ratio: 0.436        Gaps: 35
  Percent Similarity: 52.681      Percent Identity: 16.404

alignment_block:
US-09-041-994-2 x US-09-335-032-12217 ..
Align seg 1/1 to: US-09-335-032-12217 from: 1 to: 784328

330 ThrProValTyrArgPheSerLeuAlaAspGlyThrIleValThrAlaG 346
      :::::::::::::::::::: ||||| ::::::::::::::::::::
526218 TCTCGGTACACAGG...GAATAGTTGATGATTCGTGCTGCTACTGTAA 526264
346 nThrLysSerLysLeuPheArgAsnProValThrAsnAspArgHisGlyP 363
      :::::::::::::::::::: ||||| ::::::::::::::::::::
526265 CACCACCAAGACAGCTGTACAGCATTAATTTCCACCATATGATTAACAT 526314
363 heValSerThrHisPheLeuGlnArgGlnGlnAsnGlyTyrArgProAsn 379
      :::::::::::::::::::: ||||| ::::::::::::::::::::
526315 TAATGATTCGCGATCCAGCGTCACATATCAAGATTCATTCGATTCAGAC 526364
380 ProAsnProValGlyGlnGlyIleArgPro.....PrometAlaG 393
      :::::::::::::::::::: ||||| ::::::::::::::::::::
526365 ATTAACAGGCGAGACATCATATGTCAACGATGATATCCACAGATTT 526414
393 yCysAsnSerSerValGlyGlyMetSerMetSerProAsnGlnGlyLeuG 410
      :::::::::::::::::::: ||||| ::::::::::::::::::::
526415 ACATTTAGACATATCGGCTGTTTCATCACTCAATTAAT..... 526454
410 lMetProSerSerArgAlaTyrGlyLeuAlaAspProSerThrThGly 426
      :::::::::::::::::::: ||||| ::::::::::::::::::::
526454 ..... 526454
427 GlnMetSerGlyAlaArgTyrGlyGlySerSerAsnIleAlaSerLeuTh 443
      :::::::::::::::::::: ||||| ::::::::::::::::::::
526455 .....AATAGTAAACATGCCCTAAATCAA 526477
443 rProGlyPro.....GlyMetGlnSerPro.....SerSerTyrGlnAla 456
      :::::::::::::::::::: ||||| ::::::::::::::::::::
526478 CCACACCCCTGTGATCATCTATCTTTCATCCCTCATCTTCATTAAGAA 526527
456 snAsnAsnTyrGlyLeuAsnMetSerSerProOHHisGlySerProGly 472
      :::::::::::::::::::: ||||| ::::::::::::::::::::
526528 ACAAGAAGAGCTCTTGTGTGCTTAACCTGCGTTGCTTCAGAT 526577
473 LeuAlaProAsnGlnGlnAsnIleMetIleSerProArgAsnArgGlySe 489
      :::::::::::::::::::: ||||| ::::::::::::::::::::
526577 GTTGAACCTCGAAGAAAGAACTGCCGCTCATCTCCAAATATATGCTTAC 526627
489 rProLysIleAlaSerHisGlnPheSerProValAlaGlyValHisSerP 506
      :::::::::::::::::::: ||||| ::::::::::::::::::::
526628 AAGTAACATTCGCTTATATCAACAGGAGATCGCGCAATATTCATGCTC 526677
506 rMetAlaSerSerGlyAsnThrGlyAsnHisSerPheSerSerSer 522
      :::::::::::::::::::: ||||| ::::::::::::::::::::
526678 CATCATCACTTCGATCTTAACGCTTTCAGAAAGCGCTGCGCTTCCTCC 526727
523 leuSerAlaLeuGlnAlaIleSerGlyValGlyThrSerLeu..... 537

```

```

526728 AATTAACGCGACCCAGCACTAGTAATATACATCGGCTCGAATACACTGCC 526777
      ::::: :::: |||:::|||||:::|||||:::
538 .....LeuSerThrLeuSerSerProGly...ProLysL 548
      |||:::|||||::: :::: |||:::
526778 ACCTCCTCTTTACCTTACCTTACCTTACCTTACCAAAATAAGCCAAATA 526827
      |||:::|||||::: :::: |||:::
548 euAspAsnSerProAsnMetAsnIleThr..... 557
      :::: |||:::|||||:::|||||:::
526828 TATATAGAGAGGCCCAATGACAGCTCATTAATTCAGAGAAATACTTTTA 526877
      :::: |||:::|||||:::|||||:::
558 .....GlnProSerLysValSerAsnGlnAspSerLysSe 569
      :::: |||:::|||||:::|||||:::
526878 GGTGAAAACCTGTAGATGATACAAAGCGGAAGATGCTCCGCGAATTC 526927
      :::: |||:::|||||:::|||||:::
569 rProLeuGlyPheThrCysAspGlnAsnProValGluSerSerMetCysG 586
      |||:::|||||:::|||||:::
526928 AACCAACACAC.....GATTAATGCTCCAGTAGCAAAATGATGGCTGC 526968
      |||:::|||||:::|||||:::
586 InsertAsnSerArgAspHisIleSerSerAspLysGluSer..... 598
      :::: |||:::|||||:::|||||:::
526969 GTATACCGGATACCTCGAAGCGAGATGAATGAATAACAAACAAATG 527018
      :::: |||:::|||||:::|||||:::
599 ...LysGluSerSerValGluGlyAlaGlnAsnGlnArgGlyProLeuGI 614
      |||:::|||||:::|||||:::
527019 AAGAAAGAAATAAAACATTATAGTGGGAAATAATGAAGTAATGATGACAC 527068
      |||:::|||||:::|||||:::
614 uSerLys..... 616
      |||:::|||||:::|||||:::
527069 CAGCAAAATATGCACTACATCTACTAAACAGCGCTTCACCGCACCTT 527118
      |||:::|||||:::|||||:::
617 ..GlyHisLysLysLeuLeuGlnLeuLeuThrCysSer..... 628
      |||:::|||||:::|||||:::
527119 TGGGCGATGACAGACAATACTACGCGCTCTCATGCTGTCTCCAGCAGT 527168
      |||:::|||||:::|||||:::
629 ...SerAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSe 644
      |||:::|||||:::|||||:::
527169 AATGCTGACAAATCACATATATATACAGAAAGAAACACGACGACACACAA 527218
      |||:::|||||:::|||||:::
644 rSerCysLysGluSerSerValSerValThrSerProSerGlyValSers 661
      :::: |||:::|||||:::|||||:::
527219 CGGCATATACAGTAATATGTCATCCAAATAAACCATGCGCATATGCACAA 527268
      :::: |||:::|||||:::|||||:::
661 eSerThrSerGlyValSerSerThrSerAsnMetHisGlySerLeu 677
      |||:::|||||:::|||||:::
527269 ATTCTAAACCCCACTTGACGCGCTTCTACCTCTTAAC...AATATGCAATA 527315
      |||:::|||||:::|||||:::
678 LeuGlnGluLysHisArgTyrIleLeuHisLysLeuGlnAsnGlyAsnSe 694
      :::: |||:::|||||:::|||||:::
527316 AATGACGACTCTCATGAG.....AGTAATTCACAGAAA 527347
      :::: |||:::|||||:::|||||:::
694 rProAlaGluValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrS 711
      |||:::|||||:::|||||:::
527348 ACCAACAAAGCGCGATTTTTCGCTCAGCGCT..... 527384
      |||:::|||||:::|||||:::
711 eSerIleLeuThrSerCysGlyAspGlyAsnValValLysGlnGlnLeu 727
      |||:::|||||:::|||||:::
527385 .....ACACTGTAGGTGAATAAGAAATTAAGTATCTGAGGGAACA 527426
      |||:::|||||:::|||||:::
728 SerProLysLysGlnAsnAsnAlaLeuLeuArgTyrLeuLeuAspAr 744
      :::: |||:::|||||:::|||||:::
527427 TTTGTTTATGAATCGGACAGTAAATTCG...ACTAAACACCTAATATTTC 527473
      :::: |||:::|||||:::|||||:::
744 gAspAspProSerAspAlaLeuSerLysGlyLeuGlnPro..... 757
      |||:::|||||:::|||||:::
527474 TGACTCTCCAGCAGCAGCAGCAGCAGCAGCAACCTCCAAACAAAC 527523
      |||:::|||||:::|||||:::
758 .....GlnValGluGlyValAspAsnLysMetSerGlnCysThr 770
      |||:::|||||:::|||||:::
527524 ACCAACAGCAACAAATAATGAAATTAACCTCAAGATA..... 527561
      |||:::|||||:::|||||:::
771 SerSerThrIleProSerSerGlnGluLysAspProLysIleLysThr 787
      |||:::|||||:::|||||:::

```

```

527562 ACCGCCCATTTGCTAATACATTAACAAATAATTTTAAGCCGACTGAAA 527611
      :::: |||:::|||||:::|||||:::
787 rGluThrSerGluGluGlySerGlyAspLeuAspAsnLeuAspAlaIleL 804
      :::: |||:::|||||:::|||||:::
527612 TTCAAGACATATATAGACATGCTGCGCATATATTGAATTAAC.....ACAA 527652
      |||:::|||||:::|||||:::
804 euGlyAspLeuThrSerSerAspPheThrAsnAsnSerIleSerSerAsn 820
      :::: |||:::|||||:::|||||:::
527653 TCGCGACTATTAAGCAACAAATCCGAACTTGAATTTTAATGTGTGCAAGC 527702
      :::: |||:::|||||:::|||||:::
821 GlySerHisIleLeuGlyThrLysGlnGlnValPheGlnGlyThrAsnSer 837
      :::: |||:::|||||:::|||||:::
527703 AATTAACATCTGATGCTGCGGACACATCACTCGACAGATTAAGCATGT 527752
      :::: |||:::|||||:::|||||:::
837 uGlyLeuLysSerSerGlnSerValGlnSerIleArgProProTyrrAsn 854
      :::: |||:::|||||:::|||||:::
527753 AAACAGAGCCACCCCATCAATTCGACGACGACCAACCCACCA..... 527795
      :::: |||:::|||||:::|||||:::
854 rGluValSerLeuAspSerProValSerValGlySerSerProProVal 870
      :::: |||:::|||||:::|||||:::
527796 .....ATGGATGTACATCGTAGATTGCTATACCTGTACAAACCCACAC 527840
      :::: |||:::|||||:::|||||:::
871 LysAsnIleSerAlaPheProMetLeuProLysGlnPro..... 883
      :::: |||:::|||||:::|||||:::
527841 ACCAATGTTATGCC.....AGTCCGCTGATTAAGAGCTC 527875
      :::: |||:::|||||:::|||||:::
884 .....MetLeuGlyLysAsnProArgMetLeuAspSerGlnG 896
      :::: |||:::|||||:::|||||:::
527876 AAGCTTAGATGCCCTTATTAAGTTCTTCACATTTACTTATCCACAT 527925
      :::: |||:::|||||:::|||||:::
896 LysAsnTyrrGlySerSerMetGlyGlyProAsnArgAsnValThrValThr 912
      :::: |||:::|||||:::|||||:::
527926 CAAGCAACGCTAACACAAATATCGTGTCCCAATGTGCCCAATTCGCGAG 527975
      :::: |||:::|||||:::|||||:::
913 GlnThrProSerSerGlyAspTyrrGlyLeuProAsnSerLysAlaGlyArg 929
      :::: |||:::|||||:::|||||:::
527976 GAATTGGAA.....CCAAACAAATGATATTTCACAC 528004
      :::: |||:::|||||:::|||||:::
929 gMetGluProMetAsnSerAsnSerMetGlyArgProGlyGlyAspTyrr 946
      :::: |||:::|||||:::|||||:::
528005 GAAGAAATCTCTTCCAAATTTACTTTGAGACATTCCTCTGCTAATAGAA 528054
      :::: |||:::|||||:::|||||:::
946 snThrSerLeuProArgProAlaLeuGlyGlySerIleProThrLeuPro 962
      :::: |||:::|||||:::|||||:::
528055 AT..... 528056
      :::: |||:::|||||:::|||||:::
963 LeuArgSerAsnSerIleProGlyAlaArgProValLeuGlnGlnGln 979
      :::: |||:::|||||:::|||||:::
528057 .....TCTAATTTATGCTGACAAACAAAGGCTCTTAGAACACAGCTGTC 528100
      :::: |||:::|||||:::|||||:::
979 nGlnMetLeuGlnMetArgProGlyGluIleProMetGlyMetGlyAla 996
      :::: |||:::|||||:::|||||:::
528101 AAGATATTTGATTCAACCCATAATGAGAGCTCTTTA..... 528137
      :::: |||:::|||||:::|||||:::
996 snProTyrrGlyGlnAlaAlaIleAsnAsnGlnLeuGlySerTyrrProAsp 1012
      :::: |||:::|||||:::|||||:::
528138 .....CGGAGATACCTGCGGTACCGGAT 528161
      :::: |||:::|||||:::|||||:::
1013 GlnMetLeuSerMetGlnGlnValSerHisGlyThrGlnLysArgProLe 1029
      :::: |||:::|||||:::|||||:::
528162 ...CAGCTTAATCTAGAAAGATTAATCATCGAACAGCCGCAATTAATTCAC 528208
      :::: |||:::|||||:::|||||:::
1029 uLeuArgAsnSerLeu.....AspAspLeuValGlyPro..... 1040
      :::: |||:::|||||:::|||||:::
528209 AATGCAAAATATAGTGTGAATAAGATGAATTTTATGAACAGCAGCAATA 528258
      :::: |||:::|||||:::|||||:::
1041 .....ProSerAsnLeuGluGlyGlnSerAspGluArgAlaLeuLeuAsp 1055
      :::: |||:::|||||:::|||||:::
528259 ACTTCCGCTATGTTAAACCTTTATGCGATGAACAAATGTTATTTGAAGAG 528308
      :::: |||:::|||||:::|||||:::
1056 GlnLeuHisThrLeuLeuSerAsnThrArgAlaThrGlnGlyGlnGlu 1072
      :::: |||:::|||||:::|||||:::
528309 GAA.....AATTAATGCTGACTGCTAATGTAATTCGA... 528341
      :::: |||:::|||||:::|||||:::

```

```

1072     expsrghlaleugylileprogllevalasnglnglylnmlaleug 1089
      ..... |||::|||
528342     ....CCGCAACACAGTACCCTTCAGCATAGTTTA 528372
      ||::|||||
1089     luProLygInAspaLaPhegInglyngInguaLaLaValMeMetasp 1105
      ||::|||||
528373     TTCGAAGAATACCAAAAGTGATAAAGCATATTCACTCCTGTTT... 528419
      ||::|||||
1106     GlnLysAlaGlyLeuTyrgLyngInThrTyrrProAlagInglyProprome 1122
      ||::|||||
528420     .....TATTATATCATMAGAACGATTTAGAAAACAAACCGCT 528457
      :::::
1122     tGlnGlygIlyphe.....HisengInglyIns 1132
      :::::
528458     AATATCCAGTTTGTGTGAAGATGAAGACGTGATGATTATGATGCCCAA 528507
      :
1132     erProSerPheasnSermetmetasnImetasngInglylnPhe 1148
      :::::|||||
528508     ANGCTACTTTCAACAGTACTACTATGGCTGCAGCATCCACACGACGACTT 528557
      :::::|||||
1149     ProLeugInglyMetHisPro...ArgAlasn..... 1158
      |||:::|||
528558     CCATTACATGGAAGATGCTTCAGAACATATATGATTACTACAGATT 528607
      :::::
1159     lleMetArgProArghThrAsnThrProLygInleuArgMetGInleug 1175
      :::::|||||
528608     TATGTTGGCAACAATACTGGCAATACACCAATTGATATATATATACC 528657
      :::::
1175     lngInArgLeu...GlnGlyngInphleuasngInSerArglnAla 1190
      ||:::|||||
528658     CCTTAAGATGAAGCTGTCCAAGACACCTATCAAGAACAAACAATAGC 528707
      :::::
1191     leu 1191
      :::
528708     ATA 528710

seq_name: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:us-08-956-171C-59

seq_documentation_block:
; Sequence 59, Application US/08956171C
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
;         Gil H. Choi
;         Patrick S. Dillon
;         Craig A. Rosen
;         Steven C. Barash
;         Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171C
FILING DATE: 20-Oct-1997
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Hoover, Kenley K.

```

```

?      REGISTRATION NUMBER: 40, 302
?      REFERENCE/DOCKET NUMBER: PB246P1
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: (301) 610-5790
?      TELEFAX: (301) 309-8439
?      INFORMATION FOR SEQ ID NO: 59:
?          SEQUENCE CHARACTERISTICS:
?              LENGTH: 31096 base pairs
?              TYPE: nucleic acid
?              STRANDEDNESS: double
?              TOPOLOGY: linear
?      SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-08-956-171C-59

alignment_scores:
    Quality: 216.00           Length: 1502
    Ratio: 0.299             Gaps: 71
    Percent Similarity: 48.069   Percent Identity: 18.242

alignment_block:
US-09-041-994-2 x US-08-956-171C-59 ..

Align seg 1/1 to: US-08-956-171C-59 from: 1 to: 31096

55 AsnLserAspIleaspasnPhesansValysProasplyscysalaIl 71
|||||.....:|||||...::
20195 AACTTACCACTTACAAAAAACAACAACTTGAGCATCAA.....GTTGA 20238
71 eLeuylsglturhvalArgGlNllearglNllelysgluglnglyLyst 88
:::|||||.....:|||||...::
20239 ACAAGCGCAAAATGTAGACAGGTGTAATGGTGTAAAGCTAAAGCTATA 20288
88 hr.....lleSerasnspasp 93
|| CGTTAAATCTCCCATGGTCGTCATTACGACAGAATGCATCAAAATATATAT 20338
20289 CCGTTAAATCTCCCATGGTCGTCATTACGACAGAATGCATCAAAATATATAT 20338
94 AspValGlnlyAla.....AspValSerSerThrGlyGln.. 105
:::::..:: |||||.....:
20339 ACGACGAAAACAAGTCAAAATTTATCTTGATGCATCTGCACGACAAAATA 20388
106 .....GlyVallleasplysasps 112
:::|||||.....:
20389 TAATTACAACTACTGCTGTAAATATGCAAAATGGTGTATTAATGCAACGA 20438
112 erLeuclYrProleuleuLeuGlnAlaleuAspGlyPheLeuPheValVal 128
:::..:::|||||.....:
20439 ACAATGCCAAATGTGATGCTAAATGCGATTAAATGCG.....ATG 20476
129 AsnAtgGluAlaAsnIleValPheValSerGluAsnValThrGlnTyrlle 145
:::::|||||..:::|||||.....:
20477 GCMAATCAAGTCATACACAAAAGAGCGTTAAATGGTGCACAAAACCT 20526
145 uGlnTrlyrsGlnGluAspLeuValAsnThreSerValTyrrasnIleLeuH 162
|.....:|||||.....:
20527 ACCTCAAGCTAAACAACAAATGCAGCAACCA.....ATTAAACAACGAC 20570
162 iSglugluspArqlysasPheLeuLyAsnLeuProllysSerThrVal 178
||:::|||||.....:|||||.....:
20571 ATGACTTAACCAAAAAACA.....AAAGATGCATTAAAAAACAACAGTT 20614
179 Asngly.....ValSerThrAsnGluProGlnAtgGlnLysse 192
||:::|||||.....:|||||.....:
20615 AACCAATGCACAACGCTGATCTGATGCACAAATTAACGTTCAACACATCGAAC 20664
192 rHsrThrPheaScySatgMetleuMetlyshyrProHisAspIleLeuG 209
:::|||||.....:|||||.....:
20665 TGAATTGAACAGTGGGATGACAGCACTTAAACA.....GCTATTGG 20705
209 LuASP.....lleasnalerPro 215
-|||.....:|||||.....:
20706 CTGATTAAACAAGAACAAAAGCAACGCGTAATTAATGTCATCTGATCA 20755

```


22109GCATTAAAGCACAAGTACACAGTGCAGACGTGT 22143
 813 Asn.....AsnSerIleSerSerAsnGlySerHisIleuGlyThrIlysg 828
 22144 ATCTGCACCAAAATGGTGTGTAACATCTGCGACGTGAATTAATTAATCTGCGA 22193
 828 InGlnValPheGlnGlyThrAsnSerIleuGlyLeuIlySerSerGlnSer 844
 22194 TGACAGCTTTAAAGCGTGCATTCGTGATTAAGCTGACAGACAAACCACTAGT 22243
 845 ValGlnSerIleArgProProTyrAsnArgAlaValSerLeuAspSerPr 861
 22244 GGTACTATGTCAATGCTGATCGCATTAACGTCAACGATATGTAAGAAA 22293
 861 oVal.....SerValGlySerSerProProValIlyAsnIle. 873
 22294 AGTTACACCTGCCCAAAATATGCTTATGTTACACCAACCAACGATTAA 22343
 874SerAlaPheProMetLeuProIlysgLInPro 883
 22344 CACGACAGATGTTTACAAATGACGACGACGATGACGATGCTAAGACG 22393
 884 MetLeuGlyGlyAsnProArgMetMetLaspSerGlnIuAsnTyrGly.. 899
 22394 CAGTTAAACGGTAAATCATTAATTTAGACTAGCGAACAACAAATGCTAACAC 22443
 900SerSerMetGlyGlyProAsnArg..... 907
 22444 TGCATTTGATGCTTAACTCTTTAAATGTCGCCAAAAGCAAAACCTTA 22493
 908AsnVal...ThrVal... 911
 22494 AAGAACAAGTGGGTCAACGACGACGCTGCCAAATGTTCAAACTGTTCGT 22543
 912ThrGlnThrProSerSerGlyAspTyrGlyLeuProAsnSer. 925
 22544 GATAATGACCAAAACATTAACACCTGCATGAAAGCTTACGAGATAGCAT 22593
 926LysAlaGlyArgMetGluProMetLaspSer 936
 22594 TGCAGATGACCAACGATTAAGCAGGTCAA.....AACTACACAG 22644
 936 snSerMetGlyArgProGlyGlyAspTyrAsnThrSerLeuProArgPro 952
 22645 ATGCAAGTCAAAACAAACAACTGACTACAC..... 22666
 953 AlaLeuGlyGlySerIleProThrLeuProLeuArgSerAsnSerIlePr 969
 22667AGTGCAGTCA 22677
 969 oGlyAlaArgProValLeuGlnGlnGlnIleMetLeuGlnMetArgP 986
 22678 TGCAGCAAAACCAATCAT..... 22696
 986 roGlyGluIleProMetGlyMetGlyAlaAsnProTyrGlyGlnAlaIa 1002
 22697GCTCAACCACTAGTCCATCAATGATGCGCAA 22729
 1003 AlaSerAsnGlnLeuGlySerTyrProAspIlyMetLeuSerMetGluI 1019
 22730 GAAATTAATCAAGC.....AAAGCAACGATGACGCTAAACACAA 22770
 1019 nValSerHisGlyThrGlnAsnArgProLeuLeuArg..... 1031
 22771 AGCGTTAAACGCTCAAGAAAC.....TTAAGAACTGCCGCAAAACA 22811
 1032AsnSerLeuAspAspLeuValGlyProProSer 1042
 22812 ATGCGAAGCAACATTTGAACGCTTAAGTGACTTACTGACGCTCAAAA 22861
 1043 Asn.....LeuGlyGly..... 1046
 22862 GATGACAGTGAACGTCAAATGCAAGGTGCAACGATGTTAATGAAGTAA 22911

1047 GlnSerAspGluArgAlaLeuLeuAspGlnLeuHisThrLeuSerA 1063
 22912 ACAGACACAAATAATATGCG.....GATGATTTAATACCTATAGCAGA 22955
 1063 snThrAspAlaThrGlyLeuGlnGluIleAspArgAlaLeuGlyIlePro 1079
 22956 ACTTG...AAAATGTTATCAAGATCAGAAATAC.....ATTAG 22993
 1080 GluLeuValAsnGlnGlyGlnAlaLeuGluProIlysgLInAspAlaPheG 1096
 22994 CAAGGTGTTAACTTCATCAGATGCCGACGACGACGAAACGTAATGATAT.. 23041
 1096 nGlyGlnGlnAlaAlaValMetMetLaspGlnLysAlaGlyLeuTyrGly 1113
 23042ACAATGCGAGTGACGACGATGAACAAATTTAAT..... 23077
 1113 InThrTyrProAlaGlnGlyProProMetGlnGlyGlyPheHisIleuGln 1129
 23078AAGCAACAGTCCA.....AATCTTCAAAACAC 23107
 1130 GlyGlnSerProSerPheAsnSerMetMetLaspGlnMetAsnGln..G 1145
 23108 GGTGTGCAAACTGCCGTGAGAAATGTACACGCTGCTAATAAAGCAATTGA 23157
 1145 nGlyAsnPheProLeuGlnGlyMetHisProArgAla...AsnIleMet 1161
 23158 CGGTATATCAAAATGTTGCGAACGCTTACAGACACTGCGAATAATGATGTA 23207
 1161 rGProArgThrAsnThrProIlysgLInLeuArgMetGlnLeuGlnArg 1177
 23208 ATTAACCTAACATCATTAATTAATGACAAAAGAAAGCATTTGAATACAA 23257
 1178 LeuGlnGlyGlnGlnPhe.....LeuAsnGln...SerArgGlnAl 1190
 23258 ATTGAAGGTGCGACAAACAGTTGACAGGTAAATCAAGGCTTACAAACGCG 23307
 1190 aLeuGluLeuLys.....MetGluAsnProThrAlaGly.....Gly 1203
 23308 ATCTGAATTAATTAATACGCAATGACCAACTTACAAATGTTAATGATG 23357
 1203 AlaAlaValMetArgProMetMetGlnProGlnGlnGlyPheLeuAsnAla 1219
 23358 AAGCGAGTCAAAACGACGCTTAATGTTACTCAAAACCTTGAAAAACCT 23407
 1220 GlnMetValAlaGlnArgSerArgGluLeuLeuSerHisIlySphenArg 1236
 23408 AAACAAACAGCGAAATACAGCAATTGACGTTAAGCCATTTAACAAATGC 23457
 1236 nGlnArgValAlaMetMetMetGlnGlnGlnGlnGlnGlnGlnGlnGln 1253
 23458 ACAAAAGAGGCAATTAATAAATGTTGTTACAAACATCGACTGTGTCAG 23507
 1253 InGlnGlnGlnGlnGlnGln..... 1260
 23508 AAGCAAGATGATATGAGCAAAAGCAACAATGTTGATGCGACGATGAC 23557
 1261GlnGlnGlnGlnGln 1266
 23558 AAATTAAGTCAAAAGTATTCAGATTAATGCGACAAACAAACAAACCAAA 23607
 1266 nGlnThrGlnAlaPheSerProProProAsnVal.....ThrAlaSerP 1281
 23608 TTATACGTGATCAAGTCAAGTAAATAAAGAGCGGACAAATATGCTGCA 23657
 1281 roSerMetAspGlyLeuLeuAlaGlyProThrMetProGlnAlaProPro 1297
 23658 CACTGCGCAAGGTAATTTATCATCAAACTACAAAGTCCAACTTTGATCG 23707
 1298 GlnGlnPheProTyrGlnProAsnTyrGlyMetGlyGlnGlnProAspPr 1314
 23708 ACTGTATC.....AATCA 23721

425988 TCCACAGTGGGAACACACCAATAGTTCGTAGATGACGGCTT... 425944
 378 OASPrOASnPrOValGlyInGlyLeaArgProPROMeTAlaGlyCysA 395
 425943TCATCATTCGGAGAGCTCACACCCCTCTTC 425913
 395 snSerValGlyGlyMetSerMetSerProAnGInGlyLeuGlnMet 411
 425912 AATCTTACATTATATGACCAATCTTACACCCAGTATATATCAAGGCTAGT 425863
 412 ProSerSerArgAlaTyGly.....Le 419
 425862 CCCCGGAAAAAGCTTCTCAGTAGAAAAATCATCCATTAGCTTTAT 425813
 419 uAlaAspProSerThrThrGlyGlnMetSerGlyAlaArgTyGlyGlys 436
 425912 AGCAAGTCCAATATGTCGGCAGACTTCAATTGATCTACTGCGACGA 425763
 436 eSerSerAlaLeuSerLeuThrProGlyMetGlnSerProSer 452
 425762 CTGCGCTGTGTGTAAGTCTACTTCAGGA.....AGAGCGCGAC 425722
 453 SerTyGlnAsnAsnAsnTyGlyLeuAsnMetSerSerProProlHisGly 469
 425721 TCTAACCAAAATTCATTTAGGAAGCTCTCAGATTACCA.....AA 425678
 469 ySerProGlyLeu.....AlaProAsnGlnInAsnI 480
 425677 CTGCGCTATCTGTGCTGTGTAATATCCATTCACCCACACAAAGA 425628
 480 lMetIleSerProArgAsnArgGlySerProLysIleAlaSerHisGln 496
 425627 TCCACATAGAACCACTA...CCGCCGTGCCAAAAGCACCTCTAAGAAGA 425581
 497 PheSerProValAla.....GlyValHisSerPrometAl 508
 425580 AGCAGTTCCTGGCGGAGTAGTTCATATTGGTAGTATGTCAGTCGC 425531
 508 aSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerLeuSerA 525
 425530 GTCA.....AGAGGTCTTCTCATTCACACGAAATCGAGCTCCCGG 425487
 525 lAlaGlnAlaIleSerGlyGlyValGlyThrSerLeuLeu...SerThr 540
 425486 ATGCAAGCGCATCTCT.....AAGCTTCTTTCGTGAGCAATACT 425446
 541 leuSerSerProGlyProLysLeuAspAsnSerProAsnMetAsnIleTh 557
 425445 CTAGATTG.....CAAGTGTATAGTAATAGT..... 425419
 557 rGlnProSerLysValSerAsnGlnAspSerLysSerProLeuGlyPheT 574
 425418T 425418
 574 yrcysAspGlnAsnProValGlySerSerMetCysGln.....Ser 587
 425417 ATGGCAGTACATTCCTTACGACCAAGATATTCAGAAATATGCAATGAC 425368
 588 AsnSerArgAspHisLeuSerAspLysGlySerLysGlnSerValGly 604
 425367 AATGGAAGAGATGAAAGAAAGATGGGACAAATCAAGACATGAGCTGAT 425318
 604 uGlyAlaGlnAsnGlnArgGlyProLeuGlnSerLysGlyHisLysLysL 621
 425317 CTGACCTTCCAATTAAGG...CCCATC..... 425293
 621 euLeuGlnLeuLeuThrCysSerSerAspArgGlyHisSerSerLeu 637
 425292TTTTCACCATTT 425281
 638 ThrAsnSerProLeuAspSerSerCysLysGlnSerSerValSerValTh 654
 425280 AATGCATTCATATTAAGCGCAATTTCAAGCAAGTGAATTT...TTTTC 425234

654 rSerProSerGlyValSerSer.....S 662
 425233 TTCAAAATCATATATATATGAAAAGAGTAGAAGCTTCTTACATTT 425184
 662 eThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeu 678
 425183 CAAGCCCTGAAGGTGTAAGAAAGTACAAATTCGAACATTAACGCCATATC 425134
 679 GlnGlyLysHisArgGlyLeuHisLysLeuGlnInAsnGlnLysAsnSerP 695
 425133 GAGTTAGAACCTTATTAACCCGAGATCATTTGATATGCTTGGCTGTCGAC 425084
 695 oAlaGlnValAlaLysIle.....ThrAlaGlnAlaThrG 707
 425083 CGCCTGTGTAGACATTTTGCAGAGTCTTTTACACGCGCAAAATTAATG 425034
 707 Lys..... 708
 425033 GCAATATGCTATTATGCTCTGTGACCTGACCTGCTTATCATGCTTAA 424984
 708 708
 424983 ATATACGAATTCATTTTATACGAAGTATACAGAAGTCTGTATTA 424934
 709AspThrSerSerIleThrSerCysGlyLysP 719
 424933 GATTAATCAGCAAAAAGAGAGAGCTTTTCAAGTACT...TGCAGACAC 424887
 719 LysAsnValValLysGlnGlnInLeuSerProLysLysGlnAsnAsn 735
 424886 TACTAAATCTCTTCCAGCAAAAGTTTCGAGATATTGCAATGAAATAGAG 424837
 736 AlAlaLeuArgTyLeuLeuAspArgAspAspPro...SerAspAlaLe 751
 424836 GCGAAATTAAGAAGAGAGTTGGATTCTAGCAATCATATCAGACACTT 424787
 751 uSerLysGlnLeuGlnProGlnValGlnGlyValAspAsnLysMetSerG 768
 424786 A.....GAGATTAACGTTCTCCAGCCAAATATGAGACMC.....GGAT 424749
 768 lncysThrSerSerThrIleProSerSerSerGlnGlyLysAspProLys 784
 424748 ACGTTGATTCATATCTTTTGGCAAAACACTTCAATGGCGCAA..... 424708
 785 lIleLysThrGlnThrSerGlnGlnGlySerGlyLysLeuAspAsnLeuAs 801
 424707 ...TCATATACACTACAGGATGACAGTAATGACAA..... 424675
 801 pAlaIleLeuGlnLysPheLeuThrSerSerAspPheTyAsnAsnSerIleS 818
 424674ACCGAAATAGATATGACACACCGGGAATTT 424644
 818 eSerAsnGlySer.....HisLeuGly 825
 424643 CTACATATAGATTCCTTCCCAATAGGCTGTCTACTGTTTGGCGGT 424594
 826 ThrLysGlnGlnValPheGlnGlyThrAsnSerLeuGlyLeuLysSerSe 842
 424593 GGTCTTAGAGACAGCCCAATTAATTCATTGTTCAATTCATATCCACAGTAA 424544
 842 rGlnSerValGlnSerIleArgProProTyAsnArgAlaValSerLeuA 859
 424543 TTGCGGTGTTCAAAATATTCGC.....CGCTGTGTGCTTAT 424506
 859 sSerProValSer..... 863
 424505 TTAGCCAGCAGCAGCAAGAAATCCAGCTTAATAGGGAACAGATCA 424456
 864ValGlySerSerProProVa 870
 424455 TTAGATACCTCGATCAATTTACTGATGATGTCTCGATATAGTACAGAGT 424406

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (35)..(3690)
US-09-654-935A-142

alignment_scores:

Quality: 215.50 Length: 1247
Ratio: 0.396 Gaps: 61
Percent Similarity: 43.625 Percent Identity: 20.529

alignment_block:

US-09-041-994-2 x US-09-654-935A-142 ..

Align seg 1/1 to: US-09-654-935A-142 from: 1 to: 3690

404 SerProAsnGlnGlyLeuGlnMetProSerSerArgAlaTyrGlyLeuAl 420
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
127 TCCTCTCCGAGGATGATCGCTATTCCTCCAGCTCAGAGTGGGGACAGCG 176
420 AAsPProSerThrThrGlyGlnMetSerGlyAlaArgTyrGlyLeuSers 437
: : : : : : : : : : : : : : : : : : : : : : : : : :
177 GGAAGTAAAGTCTGGCTCA..TCACATGCAAGATAAATCCACAG 223
437 eRAsnIleAlaSerLeuThrProGlyProGly..MetGlnSerProSer 452
: : : : : : : : : : : : : : : : : : : : : : : : : :
224 ACCACCTCATCTCCAGCATGCTGTGAAAGGTGATCTCATGCTCCCGT 273
453 SerTyrGlnAsnAsnAsnTyrGly.....LeuAsnMetSerSerProPr 467
: : : : : : : : : : : : : : : : : : : : : : : : : :
274 CGCACATGAGATGAGGGCATGCCACCATCTTTCACAGAGACTTAGACC 323
467 oHisGlySerProGly.....LeuAlaProAsnGlnGlnAsnI 480
||| ||||| ||||| ||||| ||||| : : : : : : : : : :
324 TCAT...TCCCTGTGTAACCCGACATGTTGGCCCTGTGATCATCAGCG 370
480 leMetIle...SerProArgAsnArg.....GlySerProLysIleAla 493
: : : : : : : : : : : : : : : : : : : : : : : : : :
371 TCCCTCTAAATGCTCCAGCAATAGGAGAAATGGTCCCAAGCTATGCCG 420
494 SerHisGlnPheSerProValAlaGlyValHisSerPromet..... 507
: : : : : : : : : : : : : : : : : : : : : : : : : :
421 TATAATTGTAGAAACAACTGCCCTTCCACGCCACCCCAAGATGTGGA 470
508AlaSerSerGlyAsnThrGlyAsnHisSerPheSerSers 522
: : : : : : : : : : : : : : : : : : : : : : : : : :
471 TGGCAAGAGTGAATCTAGTTATTCAGGGGGGAGGACAGACAGCTCGG 520
522 eRLeuSerAlaLeuGlnAlaIleSerGlnGlyValGlyThrSerLeuLeu 538
: : : : : : : : : : : : : : : : : : : : : : : : : :
521 AGCCCTGGGAATACAAATCTCAGAGTAATGGAAGGGCATCCCTCTAGAG 570
539 SerThrLeuSerSerProGlyProLysLeuAspAsnSerProAsnMetAs 555
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
571 CCGGATTTAGCACTCTGGCTATTCCATCCCAAACT...AACATAG 617
555 nIleThrGlnProSerLysValSerAsnGlnAspSerLysSerProLeu. 571
: : : : : : : : : : : : : : : : : : : : : : : : : :
618 CAGCTGAGTTTGGACCAAACTCCCAACAAAGAGATCTGGTCGCTGT 667
572GlyPheTyrCys.....AspGlnAsn 578
||| ||||| ||||| ||||| : : : : : : : : : : : : : :
668 ATTCTGAGGACCAAGATGCTACTGTGATCTGTGCACATGACTCTGGA 717
579 ProValGlnSerSerMetCysGlnSerAsn..... 588
: : : : : : : : : : : : : : : : : : : : : : : : : :
718 CATGATCTGGGAATCTGTGCAATAGCAGTATGGCTTTGGGAACCCAG 767
589S 589
768 GCACAGCGTGATCAATGTTTTTTGTGGAAGAGCTCAGAAAAACCAAGGG 817

589 eArgAspHisLeuSerAspLysGlu...SerLysGlnSerSerValGlu 604
: : : : : : : : : : : : : : : : : : : : : : : : : :
818 ACCGCTCCAAATACCAGATAAATCTTATCAAGAAATCTCTTTGAG 867
605 GlyAlaGlnAsnGlnArgGlyProLeuGlnSerLysGlyHisLysLe 621
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
868 AAAGCA.....AGAACCTCCCTGCCACCTCCCGGACAGACTCCCT 911
621 uLeuGlnLeuLeuThrCysSerSerAspArgGly..... 633
: : : : : : : : : : : : : : : : : : : : : : : : : :
912 CCGCAGATCTCCAAAGAGACAGCAGCTGCAAGCGGAGGTGCTCAAG 961
634HisSer.....SerLeuThr... 638
||| ||||| ||||| ||||| : : : : : : : : : : : : : :
962 AGACCTGATGCCACACTCCAGCACTGTGTCAGACTGAGCTCCAGGC 1011
639AsnSerProLeuAspSerSerCysLysGlu..... 648
: : : : : : : : : : : : : : : : : : : : : : : : : :
1012 AAAGTGGCAGCTGCGCTCCACAGAGCCCTGCACTGATGTAAGAGCC 1061
649SerSerValSerValThrSerProSerGly 659
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1062 CTGGCTGCCCGCTCCCGAGCCAGACAGCAGTATGCTGCGACAGCA 1111
659 aLserSerSerThrSerGlyGlyAla..... 667
: : : : : : : : : : : : : : : : : : : : : : : : : :
1112 TGACTTCGCGCACCAACCCCAATGCTATCTCTGTGCGGGGCCAGCCA 1161
668SerSerThrSerAsnMetHisGlySerLeu..... 677
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1162 TCGCAGAGTGCACACAGAGCGCTCAAGTCAGATACAGGACCCCTGGG 1211
678LeuGlnGluLysHisArgIleLeuHisL 687
: : : : : : : : : : : : : : : : : : : : : : : : : :
1212 TTATTACATTGACTACAGCGCATGACAGAA..... 1242
687 yLseLeuGlnAsnGlyAsnSerProAlaGluValAlaLysIleThrAla 703
: : : : : : : : : : : : : : : : : : : : : : : : : :
1243GATCCGGGGAAC..... 1254
704 GlnAlaThrGlyLysAspThrSerSerIleThrSerCysGlyAspGlyAs 720
||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1255 CCGGAGGGGGCTGTTCACACAGCAGTGGGTGCCCACTGGGAACGGGCC 1304
720 nVal..... 721
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1305 AGTCCGCCATGTCCAAAGAGGTCCAGACCAATGCCCAAGTGCCTCG 1354
722ValLysGlnGlnLeuLeuSerProLysLysGlnAsn... 734
: : : : : : : : : : : : : : : : : : : : : : : : : :
1355 GTGGTTCAGTCAACCAAAATCATGTATCCAGAGAAATGTACACAGACTC 1404
735AsnAlaLe 737
1405 ATTCTCATCCAGTGGTATTCCAGCCAGTCGATACACCCACAGCACT 1454
737 u.....LeuArgTyrLeuLeuAspArgAspAspProSerAspAlaLeuS 752
: : : : : : : : : : : : : : : : : : : : : : : : : :
1455 CACCCCTGTGCTGTGTTTAAATCATGTCTCACCAACCAACGGGAAG 1504
752 eRlyGlnLeuGlnProGlnValGlnGlyValAspAsnLysMetSerGln 768
||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1505 GGAAG.....CCCAAGCCCAAGGTACCAAGAAAGAGATCTCTCTGTG 1545
769 CysThrSerSerThrIleProSerSerSerGlnLysAspProLysI 785
: : : : : : : : : : : : : : : : : : : : : : : : : :
1546 ATACTTTCAGATCATCTTCTCATCTGCC.....ACTTCTCT 1583
785 eLysThrGlnThrSerGlnGlnGlySerGlyAspLeuAspAsnLeuAspA 802
: : : : : : : : : : : : : : : : : : : : : : : : : :
1584 TTCTTCTAGTACTTCTACTGAAGAGAGTGCACATATGAAGAGCTGATC 1633


```

236727 TGTAAATAGACACCGTTGACAGCTTAATATTAAGTATCAAAATATT 236776
1208 ..... PromethtetGlnPro 1212
236777 GGGAAAGATATGAGAAATTCAAATTTTAAGAAAGACAAAGCTATTTTAAGACAA 236826
1213 GlnGlnGlyPheLeuAsn..... 1218
236827 CAACAGCAAAATGGCAACAAACAAATGGGAACCCAGCAGCTACTTCTACTG 236876
1219 ..... AlaglnMetValAlaGlnArgSerArgLul 1229
236877 AAACATATATATATTGCAACATCAGCAAAATATGCAACAGTACATACAGC 236926
1229 euLeuSerHisHisPheArgGlnArgValAlaMetMetMetGlnGln 1245
236927 AATGCGAGCAT.....TTACAGCAA.....TTGAAATATGACGACG 236961
1246 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1262
236962 CAACAAACAAACAGCAGCAACAAACAGCAGCAACAAACAGCAGCAGCAGCAGC 237011
1262 ngGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1279
237012 ACAACAGCAACAGCAGCATATATCCCTCGACTCT..... 237048
1279 lAsnProSerMetAspGlyLeuLeuAlaGlyProThrmProGlnAla 1295
237049 .....GGTGTGGCTAATTAATTCGCAATGCGTAATGCA 237081
1296 ProGlnGlnPheProGlyArgProAsnGlyGlyMetGlyGlnGln 1312
237082 CCGCGTAACAAATATCCATATATG..AACCAATAAAATACCTTAGCAT 237128
1312 oAspProAlaPheGlyArgValSerSerProAsnAlaMetMetSers 1329
237129 GGAT...TTTTGAACCTATGGAATAATACCAAAAGTTCCCGTATCG 237125
1329 eArgMetGlyProSerGlnAsnProMetMetGln...HisProGlnAla 1344
237176 CCGCGGCAACCCCATCCTCAACAGACATCAACGTAAGGTGAATGCG 237225
1345 AlAsnGlyLeuGlnSerSerGlnMetGlyGlyProSerGlyAsn... 1360
237226 AGACAAATCTATATACGATACCTGTACAGCATTCATCAACAATATA 237275
1361 ...LeuAlaArgAsnSerSerPheSerGlnGlnPheAlaHisGlnG 1376
237276 GAACATATCAATTTCAATATGCGGTAGTCAACACCACTCCAGGTCTG 237325
1376 lYAsnProAlaValGlySerMetValHisMetAsnGlySer..... 1389
237326 CATCAAAATACCGCTAAGTCAACCCCAAAATCAAAATCTTCCACTGAA 237375
1390 ...SerGlyHisMetGlyGlnMetAsnMetAsnProMetProMetSerG 1405
237376 ACCCAAACTAAATGGCACACCAACCCCAATATATGAAGCTGTACA 237425
1405 yMetProMetGlyProAspGlnLysTyr 1414
237426 GTCCTCTATGGGTGCACAAACATCATAT 237453
seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-813-206-703
seq_documentation_block:
: Sequence 703, Application US/09813206
: GENERAL INFORMATION:
: APPLICANT: Geating, David P.
: APPLICANT: Holtzman, Douglas A.
: TITLE OF INVENTION: NOCLETIC ACID MOLECULES DERIVED FROM A
: FILE REFERENCE: 1600.1023-002
: CURRENT APPLICATION NUMBER: US/09/813.206

```

```

: CURRENT FILING DATE: 2001-03-20
: PRIOR APPLICATION NUMBER: 60/090,181
: PRIOR FILING DATE: 1998-06-22
: PRIOR APPLICATION NUMBER: 09/338,425
: PRIOR FILING DATE: 1999-06-22
: NUMBER OF SEQ ID NOS: 934
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 703
: LENGTH: 6651
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(6651)
: OTHER INFORMATION: n = A,T,C or G
US-09-813-206-703

alignment_scores:
Quality: 210.50 Length: 1233
Ratio: 0.391 Gaps: 66
Percent Similarity: 43.633 Percent Identity: 21.087

alignment_block:
US-09-041-994-2 x US-09-813-206-703/rev ..

Align seg 1/1 to reverse of: US-09-813-206-703 from: 1 to: 6651
370 GlnArgGlnGlnAsnGlyTyrArgProAsnProAsnProValGlyGln 1 386
3967 AGAAGTTCACAGAGGAGGACCTTTGTCACCAAGAGAGCCAGAGAGG 3918
386 yLeaArgProPomeTalaGlyCysAsnSerSerValGlyGlyMetSer 403
3917 CAAGGCGGACCCCTGGAAGC.....CAAA 3892
403 eSerProAsnGlnGlyLeuGlnMetProSer.SerArgAlaTyrGly 419
3891 AGAAGCCACAGGTGACCTTTATCTGTGTGCGCCCTGTGGCT 3842
419 uAlaAspProSerThrThr.....GlyGlnMetSerGlyAlaArgTyrG 434
3841 GTCCTACACCTTTGTCACACAGGAGGCGCTGGGCGGCGGCGGCGG 3792
434 lYLeuSerSerAsnLeuAlaSerLeuThrProGlyProGlyMetGlnSer 450
3791 AACGGCGGCGCAAGGCGCCCAAGAGACCGGTGCGGACCAAGCAAG 3742
451 ProSerSerTyrGlnAsnAsnAsnTyr.....GlyLeuAsnMe 463
3741 CCAGGTCACACCAAGGATCAACCCACTCTTTGCCAGAGGCGCAACG 3692
463 tSerSerProOHHisGlySer..... 470
3691 GGGCGAGGGGCCAAGCAAGAAAGCAAGCAAGCAAGCAAGGAGGTC 3642
470 ..... 470
3641 GGGGCGGTCCTCACACAGGATCAACCTTTGGACAGCAAGAACCA 3592
471 .....ProGlyLeuAlaProAsnGlnGlnAsnGlyMetLysSer.Pro 485
3591 TCTCTTCACAGGGGAAGCTTCGCGCACAGGAGGATCCCTTCACAGCCA 3542
485 rg...AsnArgLysSerProLysile..... 492
3541 GATTCACAGGGGCTCCACCAATTCACAGGGGCGCCATGGAACCAAGG 3492
493 .....AlaSerHisGlnPheSerPro.....ValAlaGly 503
3491 GGACCAAGGTTCACCAACTGCTCCAGAAAGGACCTTTGGCCAGG 3442
503 AlHisSer.....Pro 506

```

```

|||||
3441 TTCACGAAAGGGCCAGGAAGACCAAGGAAACCTCTCTTACCTTCT 3392
507 MetAlaSerSerGlyAsnThrGlyAsnHisSerPheSerSerSerSer 523
      ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|
3391 CTGACCAAGGACGCGGACCAACCAACGCGCTGCTCCAGCATACCT 3342
523 uSerAlaLeuGlnAlaIleSerGlyValGlyThrSerLeu...LeuS 539
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
3341 TAAAGCCCGGAGTACCAAGCAACCAAGGACCACTCAGCACCCAGG 3292
539 eThrIleuSerSerPro.....GlyProIleu 548
      ::::::::::::::|::|::|::|::|::|::|::|::|::|
3291 GATCTCTTTCTTCCAGCAAGGCGCCAGGGGACCAAGGACCAAACT 3242
549 AspAsnSerProAsnMetAsnIleThrGlnProSerIleValSerAsn 565
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
3241 .....TCACCA.....AGACGCTCCAGCA 3222
565 nAspSerIleSerProLeuGlyPheTyrcysAspGlnAsnProValIus 582
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
3221 GGGCCAGCTCTTCCACGCGGGGACCTTTTGC.....C 3190
582 eSerMetCysGlnSerAsnSerArgAspHisLeuSerAspIleGluSer 598
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
3189 GCCCTCTTTGCCAGCCAGMACCCAGAGGGCCAGGGGCTCCAGCATTT 3140
599 LysGluSerSerValGlyLy.....AlaGluAsnGlnAr 610
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
3139 CCCAGAGGGGCCAGAAAGGACCGGANTTTCCGCCAGCAGCACCAAGGA 3090
610 gGlyProLeuGluSerIleGlyHisIleLysLeu.....621
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
3089 AACCCAGATGACATCCAGGGGACCAAGCGGTTCCGCCAGACACCTTT 3040
622 .....LeuGlnIleuThrCysSer.....628
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
3039 GGCTCCAGGACCCCAAAACATTACCAATGGGGCCAGGGGCTCCATTC 2990
629 .....SerAspArgGly.....HisSerIleuThrAsnSe 640
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2989 GGGTCCCGGCAAGGCGCAAGGGGACCAAGCATCGCTTTTAAACACCAA 2940
640 rProLeuAspSerSerCysIleuSerSerValIleThrSerProS 657
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2939 GCATCACCAAGTGTCTTACCAACCAAGGTTGGCGTCCAGACACAG 2890
657 eArgIleValSerSerThrSerGlyGlyValSerSerThrSer...As 672
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2889 GGGGGGCGCAGAAAGCCAGCAGGGCGCGGAGTAAACCAAGCTTACCAAG 2840
672 nMetHisGlySerLeuLeuGlnGlyHisArgIleLeuHisLysLeuL 689
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2839 GGTCTCCGGGGGG.....CACCAACAAA 2817
689 eugIlnAsnGlyAsnSerProAlaGluValAlaLysIleThrAlaGlnAla 705
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2816 GCTCCAGTGGGACCAAGGCGCGCTGG.....2785
706 ThrGlyLysAspThrSerSerIleThrSerCysGlyAspGlnValVa 722
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2784 .....ACCAACTTTTTCACCTTTGTTCACCAAGGGCAC.....2752
722 LysGlnGlnGlnIleuSerProLysLysGlyLysAsnAsnAlaLeuLeuA 739
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2751 .....CAAGCAGGCGCAAGAGAGACCAATGGGCGCGGTCCAGA 2712
739 rGlyr.....LeuLeuAspArgAspAspProSerAspAlaLeuSer 752
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2711 CCCAGGGGACGCCAATCTTTTCCAGGAGAAAGCATCAGCGACCTTTGG 2662
753 LysGluLeuGlnProGlnValGlyValAspAsnLysMetSerGlnCy 769
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

2661 GACCACATACCTTTTGTACCTTTTAGCCCTTGAAGACCAAGCGTG 2612
769 s.....ThrSerSerThrIleProSerSerSerGlnLysAspProL 784
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2611 CCACACGAGTTTACCAAGCATTTCCCTGAAGGCCCAAGGGGAG...CCCT 2565
784 yIleIleThrGlnThrSerGluGlnGlySerGlyAspLeuAspAsnLeu 800
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2564 GCGTTACCGGAGCTCCAGGGGCGCACCAAGCATCCCTTAGCACCAT 2515
801 .....AspAlaIleLeuGlyAspLeuThrSerSerAspPheTyrcys 814
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2514 CGTTCGGGGAAGCAACCGTTGGCCCTCGGGGACCAAGACCAAGCAGG 2465
814 nAsnSerIleSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValP 831
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2464 GGAACCTTGACACCAAGCGCTTG.....GCCAGGAAAAACCTCT 2424
831 heGlnGlyThrAsnSerLeuGlyLeuLysSerSerSerGlnSerValGlnSer 847
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2423 TCTGCGCTCTTCTCCAGAGGGGCGACAGGGCACAAGTTCTCCAAAG 2374
848 IleArgProProTyArgAsnArg.....AlaValSerLeuAspSerPr 861
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2373 AACCAACCCCTTGTACCAAGGTTTTCCTGCTTACCTGGAAGAGACC 2324
861 oValSerValGlySerSerProValLysAsnIleSer.....874
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2323 AGCAAGACCAAGGGAAGACCTTGGAATCCGGGGGAGCAGCAGCAAGGCG 2274
875 ..AlaPheProMetLeuProLysGlnProMetLeuGlnGlyLysAsnProArg 890
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2273 CTGTTCACCTCTCTTCCCAAGCGGAGCAAGCAGGCG.....2236
891 MetMetAspSerGlnGluAsnTyrglySerSerMetGlyGlyProAsnAr 907
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2235 .....CAAGGGGTCC.....2224
907 gAsnValThrValThrGlnThrProSerSerGlyAspTrpGlyLeuProA 924
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2223 .....TTAAGCTTCAAGCCT.....CTCCAT 2202
924 snSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMetGlyAr 940
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2201 CTTGGCAAGCAGGACCAAGCGCCA.....GGGG 2170
940 gProGlyLysPtyArgAsnThrSerLeuProArgProAlaLeuGlyLys 957
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2169 TCGGGAAAAAACCTTGGGCTTCCAAAGCCTTGGCG.....GGGC 2129
957 eTlleProThrLeuProLeuArgSerAsnSerIleProGlyAlaArgPro 973
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2128 TTCTCCAGCAGGCACTTTAGTCCAGGGAATCC.....2091
974 ValLeuGlnGlnGlnGlnMetLeuGlnMetArgProGlyGluIlePr 990
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2090 .....ATCACCAACAAGCT.....2076
990 oMetGlyMetGlyAlaAsnProTyrglyGlnAlaAlaIleAsnGlnL 1007
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2075 .....GAACCAAGGGCAC.....2061
1007 eugIlySerTrpProAspGlyMetLeuSerMetGluGlnValSerHisGly 1023
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2060 .....AGTGGGCTGGGGGTTCCGGGCGGACCATTTGACCGCGGG 2016
1024 ThrGlnAsnArgProLeuLeuArgAsnSerLeuAspIleuValGlyPr 1040
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
2015 GACCAGGGGCGCAGTTTGG.....1995
1040 oProSerAsnLeuGlnGlyGlnSerAspGluArgAlaLeuLeuAspLnt 1057
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1994 ..CCATCA.....GGACCAAGAAATGCCAGGCT.....1968

```



```

1057 euH1stHrLeuSerAsnThrAspAlaThrGlyLeuGluL1Leasp 1073
1067 .....TCCAAATTAAACCCCTTGGCCACCAG..... 1942
1074 ArgAlaLeuGlyLeProGluLeuValAsnGlnGlnAlaLeuGluPr 1090
1084 .....GGCAGACCAAGCTTCCACCGGAGGACCAAGCTTCCAC.. 1903
1090 GlnGlnAspAlaPheGlnGlnGlnGlnAlaValMetMetaspGlnL 1107
1092 ..CAAGAAGATCCCTTTGGGGCCAGAGGGC..... 1873
1107 ysaIaGlyLeuTyrGlyInThrTyrProAlaGlnGlyProPrometGln 1123
1172 .....CAA 1870
1124 GlyGlyPheH1sLeuGlnGlyGlnSerProSerPheAsnSerMetMetas 1140
1169 GGAGAAACCCAGCGTTCCACACGCGGAGCCCTTTGGGACCAAGCAAC 1820
1140 nGlnMetAsnGlnGlnGlnAsnPheProLeuGlnGlyMetH1sProAla 1157
1189 CACCATCTGCTCCAGGGAAAC.....CCACGGG 1791
1157 laaSnIleMetArgProArgPheAsnThrProLysGlnLeuArgMetGln 1173
1190 CT.....ACCAAGTTCCACCAAGCGCTCCGCGAGGG 1759
1174 LeuGlnGlnArgLeuGlnGlyGlnPheLeuAsnGlnSerArgGlnAl 1190
1175 GTCCGGGACAGCCCAAGTGGTCCGGGGTTCACTCGAGCTCTTGGCTTT 1709
1190 aLeuGlnLeuLysMetGlu.....AsnProThrAlaGlyG 1202
1708 CCTTCCTCTCCAGCAGGAGCCAGGGGGCTCTTTGAACCAACCAAGGGCC 1659
1202 LyAlaAlaValMetArgPrometMet.....GlnProGlnGln 1214
1658 AGGCTCTCCCTTAGCACCAAGTGTCTCTTTCATGCCAGACACCA 1609
1215 GlyPhe..... 1216
1608 GGTCTCTCCGCTGTACCCTTGGGAACAGAGGGCCCGCGGCTTGG 1559
1217 .....LeuAsnAlaGlnMetValAlaGlnArgSerArgGluL 1229
1558 GGTCCAAAGGGGCGCTCGGCACCAAGCAAGCAAGCAAGCAAT 1509
1229 euleuSerH1sH1sPheArgGlnGlnArgValAlaMetMetMetGlnGln 1245
1508 ACCGAGGAGCACCATTTGGACCTTTAGCAACCAAGCTGTCATCAGCAC 1459
1246 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1262
1458 CAGGGTTTCCAGCAGGCGCCAGCAGCAGCAGGCGGCGGCGCAGG 1409
1262 lngGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1274
1408 CTCACCCAGCGCACACCCCTGGGAGCTTCAAGAGCTCGGGGGCCCTTG 1359
1275 .....ProAsnValThrAlaSerProSerMe 1283
1358 GGGAGCAGGCTTACCCCTTTAGCACCAACAGCAACCCAGGAGGAGGAG 1309
1283 taspGlyLeuLeuAlaGlyProThrMetProGlnAlaProGln... 1298
1308 GGACCA.....GGGGGGCGGTGGGAGCAGAGGGGGCGGAGCAGCA 1266
1299 .....GlnPheProTyrGlnProAsnTyrGlyMetGlyGlnGln 1311
1265 AGTAGCACTTATCATTTCCAGAGACCAACAGAGGGCGGAGGG..... 1222

```

```

1312 ProAspProAlaPheGlyArgValSerSerProProAsnAlaMetMetse 1328
1221 .....GTCACAGGGGGGACACTCTCTCTCCTACACAGGAGGAGGAGG 1181
1328 rSerArgMetGlyProSer..... 1334
1180 G.....GGCCCATCTTGACCAAGAGGCTCCATTTTCCACAGGAGG 1140
1335 .....GlnAsnProMetMetGlnH1s..ProGlnAla 1344
1139 TGCCAGAGTTCACCTTAGGACCAAGCAGCAGCAACATCTTCTTGGA 1090
1345 AlaSerIleTyrGlnSerSerGlu.....MetLysG 1355
1089 CCATCC.....AAACCACTGAACACCTTGTGCTCCCTTTCATTCACAG 1046
1355 yTrpProSerGlyAsnLeuAlaArgAsnSerPhe...SerGlnGln 1371
1045 GAGGCAAGCTGTTCGCGCAATCTCGAGCAGCAGCTGAGGCGCAGAGGC 996
1371 lnpheAlaH1sGln.....GlyAsnProAlaValTyrSerMetVal 1384
995 CCACGGCTCACACAGCAGCAGCAGGTTTTCAGCTTCCCTCCATCATCTC 946
1385 H1sMet.....AsnGlySerSerGlyH1sMetGlyGlnMetMetas 1399
945 CATCTTTTCCAGGGGAGCACTGGGGGAGCTCGGGA..... 910
1399 nPrometPrometSerGlyMetPrometGlyProasp 1411
909 .....CCCATGGAGCTCGAA 895

```

```
seq_name: /cgn2_6/prodata/2/pna/US09_NEW_COMB.seq:us-09-543-679A-3003
```

```

seq_documentation_block:
; Sequence 3003, Application us/09543679A
; GENERAL INFORMATION:
; APPLICANT: NYCE, Jonathan W.
; TITLE OF INVENTION: LOW ADENOSINE AMTI-SENSE OLIGONUCLEOTIDE,
; COMPOSITIONS, KIT & METHOD FOR TREATMENT
; OF AIRWAY DISORDERS ASSOCIATED WITH
; BRONCHOCONSTRICION, LUNG INFLAMMATION,
; NUMBER OF SEQUENCES: 3111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: EPIGENESIS PHARMACEUTICALS, INC.
; STREET: 7 Clarke Drive
; CITY: Cranbury
; STATE: NJ
; COUNTRY: USA
; ZIP: 08512
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: N/A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/543,679A
; FILING DATE: 13-Apr-2000
; CLASSIFICATION: UNKNOWN
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/127,958
; FILING DATE: 1998-08-03
; ATTORNEY/AGENT INFORMATION:
; NAME: Amzel, Viviana
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: EPI-0067191b
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-409-3035
; TELEFAX: 413-254-9245
; TELEX: <unknown>
; INFORMATION FOR SEQ ID NO: 3003:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35459 base pairs

```



```

4885 CCCGACCCCAAGAGCGCCAGCGCCAGAGGAAACCCAGCCCC 4836
      |||::|||:: |||:: |||::
851 rOTyAsnArGAlaValSerLeuAspSerProValSerValGlySer 867
      ||:: |||::|||:: |||::
4835 CAGAGGGG.....CCCAACCCCAAGCGCCAGAAACAG 4801
868 ProProValYAsnIle.....SerAlaPheProMetLeuProly 881
      |||||:: |||::|||:: |||::
4800 CCCCCGAGCTCGGGTGGCCCCGGGGTGTGCGGACAGAGGCCACC 4751
881 sGlnProMetLeuGlyGlyAsnProArGmetMetAspSerGlnIuAsp 898
      |||:: |||::|||:: |||::|||::
4750 CACCCCGGCC.....CCCGACGACGAGCGCACAGCAGCGCC 4713
898 yrgLy..... 899
4712 CAGCGCAGCAGAGAVGGCCGVCAGCCCAAGACCCAGAAAGAGCAA 4663
900 SerSerMetGlyGlyProAsnArGAsnValThrValThrGlnThrProse 916
      ::::: |||:: |||:: |||:: |||::
4662 AACGCCAGAGCAGCCCGCGCGGCGCACGAGAGGGAGCAGG...CCGGC 4616
916 rSerGlyAspTrpGlyLeuProAsnSerIysAlaGlyArGmetGluProM 933
      ::|||:: |||:: |||:: |||::
4615 GACGGCGCAGAGAGG.....CGACACGAGCGGGGCCACAGACAACCCC 4572
933 eLAsn.....SerAsnSer.MetGlyArGProGlyGlyAspTyrAs 946
      ::::: |||:: |||:: |||:: |||::
4571 AAGACCGCAGAACGCCAGCAGCGCGGGGCAAGCC...GGAAGACCCCG 4525
946 nThrSerLeuProArGProAlaLeuGlyGlySerIleProThrLeuPro 963
      ::|||:: |||:: |||:: |||::
4524 CACCCAGACCCA.....CCCA 4508
963 euArSerAsnSerIleProGlyAlaArGProValLeuGlnGlnGln 979
      |||:: |||:: |||:: |||::
4507 GGAGAGAGCAGCAGAGAGCC.....AGGACAGCAGAG 4476
980 GlnMet.....LeuGlnMetArGProGlyGlyIleProMetG 992
      ::::: |||:: |||:: |||:: |||::
4475 GAGCGCGGAGCCCGCGCGAGAGAGGCGCCGCGAGGA...CCCGCGG 4429
992 yMetGlyAlaAsnProTyrGlyGlnAlaAlaAlaSerAsnGlnLeuGly 1008
      |||:: |||:: |||:: |||::
4428 CGCCACGCGCGCCCAACCGCGCGCGCCAGCAGCGGGGACCACAGGAC 4379
1009 ..SerTrpProAspGlyMetLeuSerMetGlnGlnValSerHisGlyThr 1024
      |||:: |||:: |||:: |||::
4378 ACCGAGAGCGCAGAGGGGCCACAGAGGCCACACAGCGGAGAGAGA 4329
1025 GlnAsnArGPro..... 1028
4328 GCGCCAGCGCGAGGCGCGGAAGCGCCAGAGCAGCCCAAGAGAGCGCC 4279
1029 .....LeuLeuArGAsnSerLeuAspSerLeuValGlyProProSerA 1043
      |||:: |||:: |||:: |||::
4278 CAAGCGCGCGCGCGCGAGACAGCGCGCCAGAGAGCGGGCCAGGGCGAC 4229
1043 snLeuGlnGlyGlnSerAspGluArGAlaLeuLeuAsp.....GlnLeu 1057
      ::|||:: |||:: |||:: |||::
4228 GAGCGGGCGGGAGCGCGCGCGCGCAAGAGAGAGAGCGCGCGGAGAAC 4179
1058 HisThrLeuLeuSerAsnThrAspAlaThrGlyLeu.....G 1070
      |||:: |||:: |||:: |||::
4178 CACAAGAGGAGAACAGAGAGCGCGCGAGACCGCGCGCGCCAGAGCCACA 4129
1070 uGlnIleAspArGAlaLeuGlyTyrIleProGluLeuValAsnGlnGlyAla 1087
      ::|||:: |||:: |||:: |||::
4128 GGAAGAGAGCGCGCGCAAGG.....CACCGGCAAC 4097
1087 laLeuGluProIysGlnAspAlaPheGlnGlyGlnIleAlaValMet 1103
      ::|||:: |||:: |||:: |||::

```

```

4096 CAGCCAGACCCCGAGCAGAAAGCCAGGGGGGCGGACCAACAGCCGCCA 4047
1104 MetAspGlnIleAlaGlyLeuTyrGly.....GlnThrTyrProAl 1117
      ::::: |||:: |||:: |||::
4046 CGCGCGGAAAGACAGCGAGCGCAGAGGAAACCGGAGCCCA... 3999
1117 aGlnIleProProMetGlnGlyGlyPheHisLeuGlnGlnGlnSerProS 1134
      ::|||:: |||:: |||:: |||::
3998 .CGCGACCCCGAGAAACGGCGGG.....GGACCCAGCCCA 3962
1134 eRPhAsnSerMetMetAsnGlnMetAsnGlnGlnIleAsnProLeu 1150
      ::|||:: |||:: |||:: |||::
3961 GAGCGCGCAGCAGG.....ACGGCGGCACCA 3936
1151 GlnGly.....MetHisProArGAlaAsnIleMetArG... 1161
      |||:: |||:: |||:: |||::
3935 ACCGGGGCAACCGCGCAGCCAGCAGCACCCCGCGCCACAGAACAGCGGC 3886
1162 .....ProArGThrAsnThrProIysGlnLeuA 1171
3885 CCCCCGCCCCCGCGCCCGCACCACCAAGAGCCAGGCGCCAGGAGCA 3836
1171 rGMetGlnLeuGlnGlnArG..LeuGlnGlyGlnGlnPheLeuAsnGlnSe 1187
      |||:: |||:: |||:: |||::
3835 GGAACCCAGCCCCCAGAGGGGCCCAAGCCCAAGCGCCCAACAGACAGC 3786
1187 r.....ArgGlnAlaLeuGlnLeu 1194
      ::|||:: |||:: |||:: |||::
3785 CCCCCGAGCTCGGGTGGCCCCGGGTCTGTGCGGAGAGGGGCCAGCC 3736
1194 yMetGlnIleAsnProThrAlaGlyAlaAlaValMetArG..... 1207
      |||:: |||:: |||:: |||::
3735 ACCCGCGCCCCCAGAGCAGAGCGCAGCGCCAGCGGCGCAGCAG 3686
1208 .....ProMetMetGlnProGlnGlnGlyPheLeuAsnAlaG 1220
      |||:: |||:: |||:: |||::
3685 AVGGCGCCVGCACACCCAGACGCCCCAGCCCA..... 3653
1220 nMetValAlaGlnArGSerArGlyLeuLeuSerHisIlePheArGlnG 1237
      |||:: |||:: |||:: |||::
3652 .....GCACCAAGAGACCCCAAGAAAGAAAGAGGAGCACCAGGAAAGC 3610
1237 lArGValAlaMetMetGlnGlnGln..... 1246
      |||:: |||:: |||:: |||::
3609 AAAAGAGAGAAAGCCAGCCACACAGAGAGACACACAGAAAGAAC 3560
1246 ..... 1246
3559 CCAAGCCAGGGAGCAGAACAGCGAGGGCCCAACCCAGCGGAGGAA 3510
1246 ..... 1246
3509 GGAGGTACAGAGAGGAGAGAGCGGCGGAACAGAGGACAGAGAGAG 3460
1247 .....GlnGln 1248
3459 AGGGAAGAGCGCCAGGCAAGACACAGACAGAGAGAGAGCCCAAGGA 3410
1249 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1265
      ::|||:: |||:: |||:: |||::
3409 AACAGACAGAGAAAGGAGGAGAAAGCCCAAGCAAGAGAGAGAGAGCC 3360
1265 nGlnGlnThrGlnAlaPheSerProProAsnValThrAla..... 1279
      ::|||:: |||:: |||:: |||::
3359 CAGGAACAAGGAAAGCCACCCACCCACCCACCCACCTGCGCAGGAAG 3310
1280 ..SerProSerMetAsp.....GlyLeuLeuAlaGly 1289
      ::|||:: |||:: |||:: |||::
3309 GAAACCCAGCAGAGAGAAAGAGCCVCGCCVCGVCGVCGVCGGGA 3260
1290 ProThrMetProGlnAlaProProGlnGlnPheProTyrGlnProAsnTy 1306
      ::|||:: |||:: |||:: |||::
3259 GVAVAAGAAACCAAGACVGAAGAAAGAA.....AACCA 3228

```

```

1306  rglYmeGlnGlyngInlPrOaSPrOblarPheGlyatGValSerSerProp 1323
1327  GCAVWVGVCAGGCGACGCVCVGVGGGAAGVAGCGVCVCCVCCVCCCC 3178
1323  roASnaA..... 1325
1317  CVCAGVCGVCVWVCVWVGVCYCCAGACGCGAAGCVGAGVAGVCAAGG 3128
1326  .....MetMetSerSerArGmetGlyProSerGlnAsnProMetGme 1339
1312  WCVWCCVCVGAATCTCGCAGVAVCAGCYCCAGCGAAGAACACACAGGG 3078
1339  tGlnHSerProGlnAlaIaIaSerIleTyrgInSerSerGluMetIleGlyT 1356
3077  CAAGAAGCCACGCGGACCCA.....CAACCGACGACACAAGAAAGGA. 3038
1356  rPrProSerGlyAsnLeuAlaIaTgAsnSer..... 1365
3037  ..AAAGCGCCACAGCGAAGAAAGAACGACCAAAAGACAAGAGCCACA 2990
1366  SerPheSerGlnGlnInPheAlaIaHisGlnGlyAsnProAlaValTygSe 1382
2989  AAGGCAACACACCGAGGAAACGCGGGGAGAAACCCACGAGAAACCCGAG 2940
1382  rMetValHisMetAsnGlyIleSerSerGlyHisMetGlyIleMetAsnMet 1399
2939  CGGGCGCCACGAGCGACGCGGACGAGCAAGAACACACAGGCGCAAGAGCCAGCG 2890
1399  snPro 1400
1389  GACC 2885

```

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12214

```

seq.documentation_block:
? Sequence 12214_Application US/09335032
? GENERAL INFORMATION:
? APPLICANT: Velculescu, Victor
? APPLICANT: Vogelstein, Bert
? APPLICANT: Knazler, Kenneth
? TITLE OF INVENTION: Characterization of the Yeast
? TITLE OF INVENTION: Transcriptome
? FILE REFERENCE: 01107.78572
? CURRENT APPLICATION NUMBER: US/09/335.032
? CURRENT FILING DATE: 1999-06-16
? PRIOR APPLICATION NUMBER: US 60/035.917
? PRIOR FILING DATE: 1997-01-23
? PRIOR APPLICATION NUMBER: US 09/012.031
? PRIOR FILING DATE: 1998-01-22
? NUMBER OF SEQ ID NOS: 12219
? SOFTWARE: SeqSeq for Windows Version 4.0
? SEQ ID NO 12214
? LENGTH: 666448
? TYPE: DNA
? ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12214

```

alignment_scores:	
Quality:	207.50
Ratio:	862
Percent Similarity:	0.549
	Gaps: 33
	Percent Identity: 19.142

alignment_block:
US-09-041-994-2 x US-09-335-032-12214/rev .

Align seg 1/1 to reverse of: US-09-335-032-12214 from: 1 to: 666448

446 ProglymetGInserProSerSertyrGlnAsnAsnTyrglyLeuAs 462
 |||::: ::::: ::|||:::~::~||| :::
 338119 CCAATCCAGAGATGATATTACATATTAAGAGTTCTAATAATAGCAATTTC 338070

462 nMetSerSerProProHisGlySerProGlyLeuAlaProAsnGlnGlnA 479
 ::::::::::::::: ||| :::::
 338069 TTTTCTCTCTCAAAAGCAT.....AACAGTAACTA 33804111
 479 snlIleMetIleSerProAsnArgGlySerProGlyIleAlaSerHis 495
 || :::: |||||::: :::
 338040 ACAATTATTACTCAAGCCAGAAATTAGACAGAG..... 338009
 496 GlnPheSerProValAlaIleAlaValHisSerProMetAlaSerSerGlyAs 512
 |||||::: ||| :::
 338008 GTACAAACACCCAGCGCTCATCA..... 337985
 512 nThrGlyAsnHisSerPheSerSerSerSerLeuSerAlaIleGlnAlaI 529
 |||||::: ||| :::::
 337984 .ACCGGAAACAT..... 337973
 529 IeSerGlnGlyValGlyThrSerLeuLeuSerThrIleuSerSerProGly 545
 ::::: ::::: |||
 337972 GTTAATCTCGCAAGGG 337955
 546 ProGlyLeuAspAsnSerProAsnMetAlaIleThrGlnProSerGlyAs 562
 ::::: ||||| ::::: ||| |||::: :::
 337954 AAGCAGGTAACCATCCAGCGCTTTTCAAAACATCACTGCGGGCAGCAGC 337905
 562 ISerAsnGlnAspSerGlySerProLeuGlyPheThrCysAspGlnAsp 579
 337904 TGTTCCTGAGTACTAAACATGACGTT.....C 337876
 579 rovalGlnSerSerMetCysGlnSerAsnSerAlaArgAspHisSerSerAsp 595
 ||| ::::: ||| ::::: |||
 337875 CTCAGATTC.....AATGAT 337859
 596 LysGlnSerGlySerSerValThrSerProSerGlyValSerSer 612
 ::::: ||||| ::::: ||| ::::: |||
 337858 AACATATATGAAGATTAGAACACAGGCGCAACAGCCAGAGAGAAAA 337809
 612 OleuGlnSerGlyHisLysLysLeuLeuGlnIleuLeuThrCysSer 629
 |||||::: ::::: |||
 337808 TCAGAAAAAGAGCAAGAAAGACCAACACAGCAG..... 337757
 629 eArAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSerSer 645
 ::::: |||||::: ::::: |||
 337774 GAAAGGCGATATAAC..... 337760
 646 CysLysGlnSerSerValSerValThrSerProSerGlyValSerSer 662
 |||||::: ::::: ||||| ::::: |||
 337759 ..AAGAAAGACACAAACAAATAGAGCAACTCTTTATCTTCAAGAA 337713
 662 rThrSerGlyValSerSerThrSerAsnMetHis.....GlyS 676
 |||||::: ::::: ||||| ::::: |||
 337712 AACACCTCTAGAGCATCTGCTTCCACCAACCAAGAAATGCTGGGCGTC 337653
 676 eLeuLeuGlnGlnLysHisArgIleLeuHisLysLeuLeuGlnAsnGly 692
 ::::: ||| ::::: ||| ::::: |||
 337662 CAATTGCTAACCAACCAAGCGGCTGTAAAGAACGAGTCTCCTCTT 337613
 693 AsnSerProAlaGlnValAlaLysIleThrAlaIleThrGlyLysAs 709
 ::::: |||||::: |||
 337612 GAAACGTTGCTGCAATTGACAAAGAAATTAACGATATTTAAGACGATGA 337563
 709 pThrSerSerIleThrSerCysGlyAspGlyAsnValValLysGlnGlnG 726
 ::::: ||||| ::::: ||| ::::: |||
 337562 CCMAAAAGCTGAAGCTAGT..... GAAGAAAAAGTTAATGAACAAGAA 337519
 726 ILeuSerProLysLysLysGlnAsnAsnAlaLeuLeuArgThrLeuLeu 742
 ::::: ||||| ::::: |||
 337518 CATTCGACCAAGAAACAGGGGAGACGCT..... 337487
 743 AspArgAspAspProCysAspAlaLeuSerGlnLeuGlnProGlnVal 759
 ::::: ||||| ::::: |||
 337486 GAACCTTTCTGAACA... AATGAAGACAGACTCCCTGAAG 337449
 759 IGlulGlyAlaAspAsnLysMetSerGlnCysThrSerSerThrIleProS 776


```

515886 .....CAACCTATTGAAAAACAGGGAACCTGATGTTTGGCAAT 515925
624 euleuThrCysSerSerAspArgGlyHisSerSerLeuThrAsnSer 640
|||
515926 TATCCACTCCCTCAACAAACGATTAAAGAAAAAGAAAGAACCTACT 515975
641 ProLeuAspSerSerCysGlyGluSerSerValSerValThrSerProse 657
|||||
515976 GAATGGAAAAATCTTACGTGACAGAGAA..... 516005
657 rGlyValSerSerSerThrSerGlyGlyValSerSerThrSerAsnMetH 674
516005 ..... 516005
674 lsgIleSerLeuLeuGlnGluValHisArgIleLeuHisLysLeuLeuGln 690
|||||
516006 .....GCACTAGTAGAAGAAAAAGAAAAAG..... 516029
691 AsnGlyAsnSerProAlaGluValAlaLysIleThrAlaGlnAlaThrG1 707
|||||
516030 .....GAAAGAGCTGAGAAAAATAGTGAAGCAAT.. 516059
707 YLysAspThrSerSerIleThrSerCysGlyAspGlyAsnValValLysG 724
516060 .....GAG 516063
724 lngIleGlnLeuSerProLysLysLysGluAsnAsnAlaLeuLeuArgTyr 740
|||||
516064 AGGAGGAAATTCACATGATACAGACGATTTAAAT.....TTAAGA... 516104
741 LeuLeuAspArgAspAspProSerAspAlaLeuSer.....LysG1 754
|||||
516105 .....CTGGCAGTGCCTAGACACTGGTTAAAGA 516136
754 uLeuGlnProGlnValGlu...GlyValAspAsnLysMetSerGlnCyst 770
|||||
516137 TCCTTAAATTGGAAAGTGAAGTGCAGTCCCAATAGACCGCTTGCGCA 516186
770 hrSerSerThrIleProSerSerSerGlnGluLysAspProLysIleLys 786
|||||
516187 CTTTCTCATCCACTTATGATGATTCCTCCCAAGCAAAAGATCATCAAT 516236
787 ThrGlnThrSerGlnGlnGlySerGlyAspLeuAspAsnLeuAspAlaI1 803
|||||
516237 ACCGAGACA.....GCTATGTATAGCTTAACAAAACCTACTGATAT 516277
803 eLeuGlyAspLeuThrSerSerAspPheTyrAsn.....AsnS 816
|||||
516278 CAATGAATACGTGACCCACACAAAGATTCCCTCCCGGAAGATATACT 516327
816 erIleSerSerAsnGlySerHisLeuGlyThrLysGlnGlnValPheGln 832
|||||
516328 CAAATTGCGTAATGCGCAATGCAATACAAAGCAAGGTCGATTTGAT 516377
833 GlyThrAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSerIleAr 849
|||||
516378 TTCATATCTAGT.....AAAAGGCTTCTTTGTTACATCCATA.. 516416
849 gProProTyrAsnArgAlaValSerLeuAspSerProValSerValLys 866
516417 .....CCGG 516420
866 erSerProProValLysAsnLysSerAlaPheProMetLeuProLysGln 882
|||||
516421 AAAGAGATCCCAAAAGAAATGCTCAGATGCTACTACATTT.....AAGCA 516464
883 ProMetLeuGlyGlyAsnProArgMetMetAspSerGlnGluAsnYrG1 899
|||||
516465 CGTGAATCCCATGCGCATGCGAGAGACTAGT..... 516497
899 ySerSerMetGlyGlyProAsnArgAsnValThrValThrGlnThrProS 916
|||||

```

```

516498 .....AATGTCAACATATCATGCAACGAGAAAAACA 516528
916 erSerGlyAspTyrGlyLeu.....ProAsnSerLysAlaGlyArgMet 930
|||||
516529 TCAACAAATGATACCTTTCTATATATATAAAAAAAATATCGGAGTGCTATTA 516578
931 GluProMetAsnSerAsnSerMetGlyArgProGlyGlyAspTyrAsnTh 947
|||||
516579 TCACAGCTTCTCATATGCTCAGCTCTCCGTAGATCGTTCATCTCCA 516628
947 rSerLeuProArg.....P 952
516629 TACCTGCCCTGTAATATGATCCGCTCAGAACTGATATTTGTAC 516678
952 rAlaLeuGlyGlySerIleProThrLeuProLeuArgSerAsnSerIle 968
|||||
516679 CGCATATGATGATGTCA.....CCCTCCAAACAAATACAGCC 516716
969 ProGlyAlaArgPro..... 973
516717 CCAGTTAGCTACGATTAAGATGCGATGATGCTTCAGAGAAAGACTTCAG 516766
974 .....ValLeuGlnGlnGlnGlnMetLeuGlnMetArgProG 987
516767 TTTGATATATCTTATAGCCAAACGCTATGACCGTTATATCTAGAGAG 516816
987 YGluIleProMetGlyMetGlyAlaAsnProTyrGlyGlnAlaAla 1003
|||||
516817 GAGACATTAAT.....TCCGCCGAGATCACT 516842
1004 SerAsnGlnLeuGlySerTyrProAspGlyMet...LeuSerMetGlnG1 1019
|||||
516843 AGTGAGAGATGATAGCTCTCAGCCAAAGATGATATACATATATGCTGAGCA 516892
1019 nValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAsp 1036
|||||
516893 ATGAAATTGATTAACAGAAACAAAGATTTAATG...AATGAGTTG... 516935
1036 sPLeuValGlyProProSerAsnLeuGlnGlnSerAspGluArgAla 1052
516935 ..... 516935
1053 LeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGlyLe 1069
|||||
516936 .....ACACTGCTGCTACAGAA..... 516953
1069 uGlnGluIleAspArgAlaLeuGlyIleProGlnLeuValAsnGlnGly 1086
|||||
516954 .....TTAGCAGAAATCGATTAAGAGGAAA 516978
1086 lAlaLeuGluProLysGlnAspAlaPheGlnGlnGlnValAlaVal 1102
|||||
516979 CTGAATTAGAAAGAAAGAAATCGGCTGATGAAACCAATTAATTCAGCACA 517028
1103 MetMetAspGlnLysAlaGlyLeuTyrGlyGlnThrLysProAlaGlnG1 1119
|||||
517029 TCTTTGACGATCTTCGAGGTGCTTTTCTGACATTTGAAAAGAA.. 517076
1119 yProProMetGlnGlyLysPheHisLeuGlnGlnSerProSerPheA 1136
517077 .....CTGAGAAAAAAATCCAGTAAATTTG 517101
1136 snSerMetMetAsnGlnMetAsnGlnGlnGlnLysAsnPheProLeuGlnGly 1152
|||||
517102 TGCATTTGATCCACGCAATTAATGACGAG.....AGATTAAAGAGA 517142
1153 MetHisProArgAlaAsnLleMetArgProArgThrAsnThrProLysG1 1169
|||||
517143 TTCATTTCGGAGGAGAGAGATTATTAATCCAAAGAAAAATGACACC...AAAC 517189
1169 nLeuArgMetGlnLeuGlnGlnArgLeuGlnGlnGlnGlnPheLeuAsnG 1186
|||||
517190 AAGTTCTATGGAATCTAGCAAGAAATTTGAATAATTAATTAATTTG 517239

```

1186 InSe:ArgGlnAlaLeuGluLeuLysMet 1195
:::||||::: |||:::|
517240 ATGAGAGAGATTCA...GAAATTGAAATG 517265


```

957 ACAAGGCAACAGGGTACCCAGTCACTTTTCCACACAGCCAGC 1006
93 eutyrglygl...ThrtYrProAlagInglyProPrometGIngly 108
   :||| :||| :||| :||| :||| :||| :||| :||| :|||
1007 AATGGCAACAGGGTCTACCCAGTCACTTTTCCAGCAGCAGCA 1056
109 pHeHsleuGInglyInserProSerPheAsnSerMetMetAsnGln.. 124
   ||||| :||| :||| :||| :||| :||| :||| :||| :|||
1057 CAAGGGCAACAGGGTACCAAGGAGTCACTTTACGACGACCGCG 1106
125 .....MetAsnGInglyInserPheProleu.... 133
   :||| :||| :||| :||| :||| :||| :||| :||| :|||
1107 ACAAGTGCACAGTGGGACAGGCAAGGCAAGGCAAGTCACTTGC 1156
134 .....GInglyMetHis 137
   :||| :||| :||| :||| :||| :||| :||| :||| :|||
1157 CTTCGCCGACGAGCCAGGACAGGCAACAGCTAGACAGGAGCA 1206
138 Pro 138
   |||
1207 CCA 1209

```

seq_name: /cgn2_6/prodata/1/lna/5b_COMB.seq:us-08-061-376-4

```

seq_documentation_block:
; Sequence 4, Application US/08061376
; Patent No. 6175000
; GENERAL INFORMATION:
; APPLICANT: Evans, Glen A.
; APPLICANT: Djibail, Malek
; APPLICANT: Selleri, Lucia
; APPLICANT: Parry, Pauline
; TITLE OF INVENTION: CHARACTERIZATION OF A CHROMOSOME 11Q23
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/061,376
; FILING DATE: 13-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9387
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)546-4737
; TELEFAX: (619)546-9392
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11907 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; US-08-061-376-4

```

alignment_scores:
 Quality: 85.50 Length: 179
 Ratio: 0.983 Gaps: 6
 Percent Similarity: 48.603 Percent Identity: 21.788

alignment_block:
 US-09-041-994-2_COPY_1018_1179 x US-08-061-376-4 ..
 Align seg 1/1 to: US-08-061-376-4 from: 1 to: 11907

```

4 SerHis.....GlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAs 18
   ||||| :||| :||| :||| :||| :||| :||| :||| :|||
6829 AGTCACCTGGATGATCTTCATCTTCAGAAATGAAGACAGTCAGCTTC 6878
18 pAsPleuValGlyProProSerAsnLeuGInglyInserAspGlnArg 35
   :||| :||| :||| :||| :||| :||| :||| :||| :|||
6879 AGACCTGGGTCCAGACGCTCTCTTAAGGACAGAGAACCAAA.... 6924
35 lAlLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGly 51
   :||| :||| :||| :||| :||| :||| :||| :||| :|||
6925 .....GTCTGAGTTCCAGAGGCTCAGAGCGA 6951
52 LeuGInGluIleAspArgAlaLeuGlyTleProGluLeuValAsnGln.. 67
   :||| :||| :||| :||| :||| :||| :||| :||| :|||
6952 TCTGCACATATGATGCTTACCTCGAATCTCTAAAGTGGCCACAGGT 7001
68 .....GlyGlnAlaL 71
   |||||
7002 TCATACACACATCTAGAGACTGATGTAGTAAATCGGCTCTTGG 7051
71 euGlnProLysGlnAspAlaPheGInglyGlnAlaAlaValMetMet 87
   ||||| :||| :||| :||| :||| :||| :||| :||| :|||
7052 CTGAACCCCTTCAGTGTCTTCTCTAAAGAGGCCCTC..... 7092
88 AspGlnLysAlaGlyLeuTyrglyInThrTyrgProAlaGInglyProBr 104
   :||| :||| :||| :||| :||| :||| :||| :||| :|||
7093 .....TCCTTCCA..... 7101
104 oMetGInglyGlyPheHisLeuGInglyGln..... 114
   :||| :||| :||| :||| :||| :||| :||| :||| :|||
7102 .....CACCTCAATTCAGAGGGCAAGAGATGTCAGACCAAC 7141
115 .....SerProSerPheAsnSerMetMetAsnGlnMetAsnGln 127
   :||| :||| :||| :||| :||| :||| :||| :||| :|||
7142 ACACGATTCACCCCAATCAGCAAACTCTCTCAGATGAAATCTGAA 7191
128 GInGlyAsnPheProLeuGInglyMetHisProArgAlaAsnIleMetArg 144
   :||| :||| :||| :||| :||| :||| :||| :||| :|||
7192 GTCAAAACCTTGACCTATCTGGAATGAGCAACAGATCTTCATATGCA 7241
144 gProArgThrAsnThrProLysGlnLeuArgMetGln 156
   :||| :||| :||| :||| :||| :||| :||| :||| :|||
7242 CGAACATATGCGATCTAGTTCAGAGATAGAGACAG 7278

```

seq_name: /cgn2_6/prodata/1/lna/5a_COMB.seq:us-08-320-559-1

```

seq_documentation_block:
; Sequence 1, Application US/08320559
; Patent No. 5631135
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods for
; TITLE OF INVENTION: Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5631135rls
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; OPERATING SYSTEM: PC-DOS

```


alignment_block:

US-09-041-994-2_COPY_1018_1179 x US-08-327-392-1 ..

Align seg 1/1 to: US-08-327-392-1 from: 1 to: 14255

```
4 SerHis.....GlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAs 18
|||||
6655 AGTCACCTTGATGATGATCTTCACAGAAATGAAGACATGCAGTGCCTTC 6704
18 PASPLEUVALGlyProProSerAsnLeuGlnGlyInSerAspGluArgA 35
|||||
6705 AGACTGGGTGCTCAGAGAGCTCCTTTAAAGGAGAGAAAGACCAAA... 6750
35 lAleuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGly 51
|||||
6751 .....GTGCTGAGTTCACAGAGCTCCAGAGGGA 6777
52 LeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnGln.. 67
|||||
6778 TCTGCACATATATGTGGCTTACCCTGGAAATTCCTAAACTGCCCCACAGGT 6827
68 .....GlyGlnAlaL 71
6828 TCATAACACACATCTAGAGACTGATGTAGTAAATCGGCTCCTTGG 6877
71 euGluProLysGlnAspAlaPheGlnGlyGlnAlaAlaValMetMet 87
|||||
6878 CTGAACCCCTTCACGTGCTTCTCTTAAAGAGGCCCTC..... 6918
88 AspGlnLysAlaGlyLeuTyrglyLeuThrTyrgProAlaGlnGlyProPr 104
|||||
6919 .....TCCTTCCCA..... 6927
104 oMetGlnGlyGlyPheHisLeuGlnGlyGln..... 114
|||||
6928 .....CACCTCCATTTAGAGGGCAAGAAGATGTCGAGACCAAC 6967
115 .....SerProSerPheAsnSerMetMetAsnGlnMetAsnGln 127
|||||
6968 ACACAGATTCTTACCCCATGAGCAAACTCCTCTCCAGATGATGATCTGAA 7017
128 GlnGlyAsnPheProLeuGlnGlyMetHisProArgAlaAsnIleMetAr 144
|||||
7018 GTCAAAACCTTGAACTATCTGGAATGAGCAACAGATCATCATTAATCA 7067
144 gProArgThrAsnThrProLysGlnLeuArgMetGln 156
|||||
7068 CGAACATATGGGATCTAGTTCAGAGATAGAGACAG 7104

seq_name: /cgn2_6/ptodata/1/lna/5A_COMB.seq:US-08-306-691B-55

seq_documentation_block:
; Sequence 55, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavoorgna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
```

CURRENT APPLICATION DATA:

```
;; APPLICATION NUMBER: US/08/306,691B
;; FILING DATE: September 15, 1994
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Monaco, Daniel A.
;; REGISTRATION NUMBER: 30,480
;; REFERENCE/DOCKET NUMBER: 8321-8
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (215) 568-8383
;; TELEFAX: (215) 568-5549
;; TELEX: No. 5734039e
;; INFORMATION FOR SEQ ID NO: 55:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 14255 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-306-691B-55
```

alignment_scores:

```
Quality: 85.50 Length: 179
Ratio: 0.983 Gaps: 6
Percent Similarity: 48.603 Percent Identity: 21.788
```

alignment_block:

US-09-041-994-2_COPY_1018_1179 x US-08-306-691B-55 ..

Align seg 1/1 to: US-08-306-691B-55 from: 1 to: 14255

```
4 SerHis.....GlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAs 18
|||||
6655 AGTCACCTTGATGATGATCTTCACAGAAATGAAGACATGCAGTGCCTTC 6704
18 PASPLEUVALGlyProProSerAsnLeuGlnGlyInSerAspGluArgA 35
|||||
6705 AGACTGGGTGCTCAGAGAGCTCCTTTAAAGGAGAGAAAGACCAAA... 6750
35 lAleuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGly 51
|||||
6751 .....GTGCTGAGTTCACAGAGCTCCAGAGGGA 6777
52 LeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnGln.. 67
|||||
6778 TCTGCACATATATGTGGCTTACCCTGGAAATTCCTAAACTGCCCCACAGGT 6827
68 .....GlyGlnAlaL 71
6828 TCATAACACACATCTAGAGACTGATGTAGTAAATCGGCTCCTTGG 6877
71 euGluProLysGlnAspAlaPheGlnGlyGlnAlaAlaValMetMet 87
|||||
6878 CTGAACCCCTTCACGTGCTTCTCTTAAAGAGGCCCTC..... 6918
88 AspGlnLysAlaGlyLeuTyrglyLeuThrTyrgProAlaGlnGlyProPr 104
|||||
6919 .....TCCTTCCCA..... 6927
104 oMetGlnGlyGlyPheHisLeuGlnGlyGln..... 114
|||||
6928 .....CACCTCCATTTAGAGGGCAAGAAGATGTCGAGACCAAC 6967
115 .....SerProSerPheAsnSerMetMetAsnGlnMetAsnGln 127
|||||
6968 ACACAGATTCTTACCCCATGAGCAAACTCCTCTCCAGATGATGATCTGAA 7017
128 GlnGlyAsnPheProLeuGlnGlyMetHisProArgAlaAsnIleMetAr 144
|||||
7018 GTCAAAACCTTGAACTATCTGGAATGAGCAACAGATCATCATTAATCA 7067
```

144 gProargThrAsnThrProlysglnLeuArgMetGln 156

7068 CGAACATATGGATCTAGTCCAGAGATGAGACAG 7104

seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:US-08-545-860D-1

seq_documentation_block:

```

; Sequence 1, Application US/08545860D
; Patent No. 6040140
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaanl, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1 Region
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESS: No. 6040140ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,860D
; FILING DATE: 07-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04496
; FILING DATE: 22-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10930
; FILING DATE: 09-DEC-1992
; APPLICATION NUMBER: US 08/327,392
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/320,559
; FILING DATE: 11-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/062,443
; FILING DATE: 14-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/971,094
; FILING DATE: 30-OCT-1992
; APPLICATION NUMBER: US 07/888,839
; FILING DATE: 27-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/805,093
; FILING DATE: 11-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca Esq., Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1262
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1425
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ANTI-SENSE: No
; US-08-545-860D-1

```

alignment_scores:
 Quality: 85.50 Length: 179
 Ratio: 0.983 Gaps: 6
 Percent Similarity: 48.603 Percent Identity: 21.788

alignment_block:

US-09-041-994-2_copy_1018_1179 x US-08-545-860D-1 ..

Align seg 1/1 to: US-08-545-860D-1 from: 1 to: 14255

```

4 SerHis.....GlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAs 18
||||| |||||.....: : : : : : : : : : : : : : : : : : :
6655 AGTCACCTTGATGATGCTTATCTTCAGAAATGAGACAGTCAGTCTTC 6704
||||| |||||.....: : : : : : : : : : : : : : : : : : :
18 PASpleValGlyProProSerAsnLeuGlnGlnGlnSerAspGlnArg 35
||||| |||||.....: : : : : : : : : : : : : : : : : : :
6705 AGACTTGCTCCAGAGCTCTCTTTAAAGGAGAGAACACCAA.... 6750
||||| |||||.....: : : : : : : : : : : : : : : : : : :
35 lalLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGly 51
||||| |||||.....: : : : : : : : : : : : : : : : : : :
6751 .....GTGCTGAGTCCAAAGAGCTCAGAGGGA 6777
||||| |||||.....: : : : : : : : : : : : : : : : : : :
52 LeuGlnGlnIleAspArgAlaLeuGlyIleProGlnLeuValAsnGln.. 67
||||| |||||.....: : : : : : : : : : : : : : : : : : :
6778 TCTGCACATATATGTGGCTTACCCCTGGAATTCCTAAACTGCCCCACAGGT 6827
||||| |||||.....: : : : : : : : : : : : : : : : : : :
68 .....GlyGlnAlaAla 71
||||| |||||.....: : : : : : : : : : : : : : : : : : :
6828 TCATACACACACATCTAGAGAACTGAATGTATTGTAATAATGGCTCTTGG 6877
||||| |||||.....: : : : : : : : : : : : : : : : : : :
71 euGlnProlysglnAspAlaPheGlnGlnGlnGlnAlaAlaValMetMet 87
||||| |||||.....: : : : : : : : : : : : : : : : : : :
6878 CTGAACCCCTTCTCAGTCTCTTTCTCTTAAGAGGCCCTC..... 6918
||||| |||||.....: : : : : : : : : : : : : : : : : : :
88 AspGlnIleAspGlnLeuThrGlyGlnThrTyProAlaGlnGlnGlyPro 104
||||| |||||.....: : : : : : : : : : : : : : : : : : :
6919 .....TCTCTCCA..... 6927
||||| |||||.....: : : : : : : : : : : : : : : : : : :
104 ometGlnGlyGlyPheHisLeuGlnGln..... 114
||||| |||||.....: : : : : : : : : : : : : : : : : : :
6928 .....CACCTCATTTGAGAGGCGCAAGAAATGATCGAGACCAAC 6967
||||| |||||.....: : : : : : : : : : : : : : : : : : :
115 .....SerProSerPheAsnSerMetLeuAsnGlnMetAsnGln 127
||||| |||||.....: : : : : : : : : : : : : : : : : : :
6968 ACACAGATCTTACCATCAGCAAACTCTCCAGATGAAGATACGAA 7017
||||| |||||.....: : : : : : : : : : : : : : : : : : :
128 GlnGlnAsnThrProLeuGlnGlnGlnMetHisProArgAlaAsnIleMet 144
||||| |||||.....: : : : : : : : : : : : : : : : : : :
7018 GTCAAAACCTTGAAGCTATCTGGAATGAGCAACAGATCAATTCATCA 7067
||||| |||||.....: : : : : : : : : : : : : : : : : : :
144 gProargThrAsnThrProlysglnLeuArgMetGln 156
||||| |||||.....: : : : : : : : : : : : : : : : : : :
7068 CGAACATATGGATCTAGTCCAGAGATGAGACAG 7104
||||| |||||.....: : : : : : : : : : : : : : : : : : :

```

seq_name: /cgn2_6/ptodata/1/lna/PCTUS_COMB.seq:PCT-US94-04496-1

seq_documentation_block:

```

; Sequence 1, Application PC/TUS9404496
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaanl, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESS: Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania

```

```
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04496
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Deluca Esq., Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-1242
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1425
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
ANTI-SENSE: NO
PCT-US94-04496-1
```

```
alignment_scores:
Quality: 85.50 Length: 179
Ratio: 0.983 Gaps: 6
Percent Similarity: 48.603 Percent Identity: 21.788
```

alignment_block:

US-09-041-994-2_COPY_1018_1179 x PCT-US94-04496-1 ..

Align seg 1/1 to: PCT-US94-04496-1 from: 1 to: 14255

```
4 SerHis.....GlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAs 18
|||||
6655 AGTCACTGGATGGATCTTCATCTTCAGAAATGAGACAGTCACGCTTC 6704
18 PASpleuValGlyProPserAsnLeuGlnGlyInSerAspGluAAs 35
|||||
6705 AGACTGGTGTCCAAGACCTCTCTTAAGGAGAGAGACCAAA.... 6750
35 lAleuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGly 51
|||||
6751 .....GTGCTGAGTTCACAGAGCTCAGAGGGA 6777
52 LeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnGln.. 67
|||||
6778 TCTGCACATATATGTGGCTTACCCTGGAATCTCTAAAGAGCCACAGT 6827
68 .....GlyGlnAlaLeu 71
6828 TCATACACACACATCTAGAGACTGAATGTAGTAAATCGCTCTCTTG 6877
71 euGluProLysGlnAspAlaPheGlnGlyGlnGluAlaValaMetMet 87
|||||
6878 CTGAACCTCTTCAGTGTGCTTCTCTAAAGAGCCCTC..... 6918
88 AspGlnLysAlaGlyLeuTyrGlyGlnThrThrProAlaGlnGlyProPr 104
|||||
6919 .....TCCTTCCA..... 6927
104 oMetGlnGlyGlyPheHisLeuGlnGlyGln..... 114
|||||
6928 .....CACCTTCATTGAGGGCAAGAGATGTCAGACCAAC 6967
115 .....SerProSerPheAsnSerMetMetAsnGlnMetAsnGln 127
|||||
6968 ACACAGATTCTACCAATCAGCAAACTCCTCTCCAGATGAGATGCTGA 7017
```

```
128 GlnGlyAsnPheProLeuGlnGlyMetHisProArgAlaAsnIleMetAr 144
|||||
7018 GTCAAAACCTTGAGCTATCTGATGATGAGCAACAGATCATCATATCA 7067
144 gProArgThrAsnThrProLysGlnLeuArgMetGln 156
|||||
7068 CGAACATATGGGATCTACTTCACAGATAGGAGACAG 7104
```

seq_name: /cgn2_6/prodata/1/lna/5A_COMB.seq:us-08-227-536-1

seq_documentation_block:

```
; Sequence 1, Application US/08227536
; Patent No. 5638784
; GENERAL INFORMATION:
; APPLICANT: Ecken, Richard
; APPLICANT: Ewen, Mark
; APPLICANT: Livingston, David
; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurigin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,536
FILING DATE: 14-APR-1994
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Williams Ph.D., Kathleen A.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: DFCI-308XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9046 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 1200..8441
US-08-227-536-1
```

```
alignment_scores:
Quality: 85.00 Length: 187
Ratio: 1.062 Gaps: 7
Percent Similarity: 42.781 Percent Identity: 23.529
```

alignment_block:

US-09-041-994-2_COPY_1018_1179 x US-08-227-536-1 ..

Align seg 1/1 to: US-08-227-536-1 from: 1 to: 9046

```
6 GlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAspAspLeuValG 22
|||||
7329 GGCACGTGTCTCAACAGCGCTTACAAACCTTTGCGGACTCTGAGGTC 7378
```

```
22 yProProSerAsnLeuGlnGlyGlnSerAspGluArgAlaLeuAspG 39
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7379 TCACAGCTCTCCCTGCACAGCAGCAGCAGCAGCAGCAGCAGCAG 7416
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
39 InLeuHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuGluIle 55
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7417 TCCTTCAC.....GCCAAGCCCGAGCTGTG.....GTCTTAGTA 7424
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
56 AspArgAlaLeuGlnGlyLeuProGluLeuValAsnGlnGlyAlaLeu 72
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7425 .....GCCAAGCCCGAGCTGTG.....GTCTTAGTA 7442
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
72 uProLysGlnAspAlaPheGlnGlyGlnGluAlaValMetMetAsp 89
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7443 .....GCTGCATTCATCAGC 7458
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
89 InLysAlaGlyLeuTyrglyGlnThrTyrrProAlaGlnGlyPro.... 103
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7459 AGCGGGCTGCAGATGATGCACTTAATCCACAAACCCATCCTGGGAG 7508
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
104 .....ProMet 105
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7509 CCTGGCATGCCCGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7558
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
105 tGlnGlyGlyPheHisLeuGlnGlyGlnSerProSerPheAsnSerMet. 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7559 TCAGCAGGGGGTCCACTCCATTCACAGCCATGCAGAGCATGAATCCAT 7608
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 .....MetAsnGlnMetAsnGlnGln 128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7609 AGCGGGGCTTCAGAGGGCTGCTGCCCGCAGCAGCAGCAGCAGCAG 7658
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
129 GlysAsnPheProLeuGlnGlyMetHisProArgAlaAsnIleMetArg 145
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7659 CTCAGCCACCCATGGAGGATGAGCCCCCAGGCTCAGCATGATGAT 7708
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
145 cArgThrAsnThr...ProLysGlnLeuArg.....MetGlnLeuGln 159
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7709 GAACCAACACACCATGCTTCACAAATTCGAGACATCTTGAGACGAC 7758
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
159 InArgLeuGln 162
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7759 AATGATGCAA 7769

seq_name: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:PCT-US95-04682-1
seq_documentation_block:
; Sequence 1, Application PC/TUS9504682
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurigin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04682
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,536
; FILING DATE: 14-April-1994
```

```
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Holliday C. Heine, Ph.D.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DECI-308xq999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9046 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1200..8441
PCT-US95-04682-1
```

```
alignment_scores:
Quality: 85.00 Length: 187
Ratio: 1.062 Gaps: 7
Percent Similarity: 42.781 Percent Identity: 23.529
```

```
alignment_block:
US-09-041-994-2_COPY_1018_1179 x PCT-US95-04682-1 ..
```

```
Align seg 1/1 to: PCT-US95-04682-1 from: 1 to: 9046
```

```
6 GlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAspAlaValG 22
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7329 GGCACGTGTCTCAACAACCTTACAAACCTTTCGGAGCTCAGTC 7378
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
22 yProProSerAsnLeuGlnGlyGlnSerAspGluArgAlaLeuAspG 39
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7379 TCACAGCTCTCCCTGCACAGCAGCAGCAGCAGCAGCAGCAGCAG 7416
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
39 InLeuHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuGluIle 55
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7417 TCCTTCAC.....GCCAAGCCCGAGCTGTG.....GTCTTAGTA 7424
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
56 AspArgAlaLeuGlnGlyLeuProGluLeuValAsnGlnGlyAlaLeu 72
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7425 .....GCCAAGCCCGAGCTGTG.....GTCTTAGTA 7442
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
72 uProLysGlnAspAlaPheGlnGlyGlnGluAlaValMetMetAsp 89
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7443 .....GCTGCATTCATCAGC 7458
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
89 InLysAlaGlyLeuTyrglyGlnThrTyrrProAlaGlnGlyPro.... 103
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7459 AGCGGGCTGCAGATGATGCACTTAATCCACAAACCCATCCTGGGAG 7508
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
104 .....ProMet 105
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7509 CCTGGCATGCCCGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7558
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
105 tGlnGlyGlyPheHisLeuGlnGlyGlnSerProSerPheAsnSerMet. 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7559 TCAGCAGGGGGTCCACTCCATTCACAGCCATGCAGAGCATGAATCCAT 7608
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 .....MetAsnGlnMetAsnGlnGln 128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7609 AGCGGGGCTTCAGAGGGCTGCTGCCCGCAGCAGCAGCAGCAGCAG 7658
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
129 GlysAsnPheProLeuGlnGlyMetHisProArgAlaAsnIleMetArg 145
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7659 CTCAGCCACCCATGGAGGATGAGCCCCCAGGCTCAGCATGATGAT 7708
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
145 oAqgThraSthr...ProlysglnLeuArg.....MetGlnLeuGln 159
      :: ||||| ||::|||::|||
7709 GAACCAACACATCGCTTCACAAATTCGAGACATCTTGAGACAGACG 7758
159 lnArgLeuGln 162
      ||::|||
7759 AAATGATGCNA 7769

seq_name: /cgn2_6/ptodata/1/lna/5A_COMB.seq:us-08-612-986-1

seq_documentation_block:
; Sequence 1, Application US/08612986
; Patent No. 5770384
; GENERAL INFORMATION:
; APPLICANT: Elliot J. Androphy
; APPLICANT: Dave E. Breiding
; TITLE OF INVENTION: E2 BINDING PROTEINS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive & Cockfield
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,986
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/361,806
; FILING DATE: 22 DEC 1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: NEP-004DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1071 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1071
; US-08-612-986-1

alignment_scores:
      Quality: 84.00      Length: 122
      Ratio: 1.647      Gaps: 3
Percent Similarity: 41.803      Percent Identity: 27.049

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-08-612-986-1 ..
Align seg 1/1 to: US-08-612-986-1 from: 1 to: 1071

73 ProlysglnaspAlaphgInglyGln...AlaAlaValMetMetas 88
      |||::|||::|||::|||
595 CCAGAGCATGACAAATTCGAGGATGCTGCTATGGAGACTGC 644
88 pGlnlysaLgIlyLeuTyrgIyInThrTyrrProAlaGlnGlyPro.... 103
```

```
645 TCAGCCCCCAGCTCATATATGGCCACAGACAGAGCTTATATGCTTAGTC 694
      |||
103 ..... 103
695 AGCAGCTCAGAGCTCCCTTCGGCATTCCTCGAGTGGAGTACATTCAG 744
104 .....PrometGlnI 107
745 CCACAGCACCAGCCCTATGCTGTGCATGGCCACTTTCAGCCCACTCAG 794
107 yGlyrPhelIsLeuGlnGlyInserProSerPhenSerMetMetasng 124
      ||||| |||::|||::|||::|||::|||::|||
795 AGGTTTCCTCAGCGCTGTGTGGCTGTCTCTGCAAAAGCAGATGGAGC 844
124 lMetasngInglyAsnPh.....ProLeuGlnGlymet 136
      :: ||||| |||::|||
845 ATGCTAACAGACAGACTGGCTTCGCGACATCATCTCTGCGCCCATG 894
137 HisProArGAlaAsnIleMetArgProArgThrAsnThrProlysglnG 153
      |||||::|||
895 CACCCCGAGGCTGCGATCCAGCCCTGAGACTCTTGTCCGCCAGCT 944
153 uArgMetGlnLeuGln 158
      |::|||::|||
945 CCTGTGAGATGCAG 960

seq_name: /cgn2_6/ptodata/1/lna/5A_COMB.seq:us-08-361-806A-1

seq_documentation_block:
; Sequence 1, Application US/08361806A
; Patent No. 5792833
; GENERAL INFORMATION:
; APPLICANT: Elliot J. Androphy
; APPLICANT: Dave E. Breiding
; TITLE OF INVENTION: E2 BINDING PROTEINS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive & Cockfield
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/361,806A
; FILING DATE: 22 DEC 1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: NEP-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1071 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1071
; US-08-361-806A-1
```



```

NAME: Shelness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: FHCRI7537
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100
TELEFAX: 206-224-0779
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1896 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: human RAD9compa cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-198-446B-8

```

```

alignment_scores:
Quality: 83.00      Length: 189
Ratio: 0.988       Gaps: 8
Percent Similarity: 44.444      Percent Identity: 25.926

```

alignment_block:

US-09-041-994-2_COPY_1018_1179 x US-08-198-446B-8/rev ..

Align seg 1/1 to reverse of: US-08-198-446B-8 from: 1 to: 1896

```

3 ValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAspAs 19
  ::::::::::::::::::::
1074 TTGTCCAGAGTCCAGAGAGAGTCCAGAGAGTCCAGAGTCCAGAGTCCAG 1025
  19 pLeuVal.....GlyProProSerAsnLeuGlnGlnGln 31
  ::::::::::::::::::::
1024 CTCGGTGGACAGATTCATCACACCCCTCTCTACTCTCTCTCTCTCTCT 975
  31 er.....AspGlnArgAla 35
  ::::::::::::::::::::
974 CCGTGGCCACAGTCCAGATTCCTCCCAATCCTCCCAACACAGAGGCGT 925
  36 LeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGlyLe 52
  ::::::::::::::::::::
924 GTTAGGCGCGAAACCTTAACTTAGTCTCGAGC..... 889
  52 uGlnGluLeuAspArgAlaLeuGlyIleProGluLeuValAsnGlnGly 69
  ::::::::::::::::::::
888 .....CAGACAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 852
  69 In..AlaLeuGluProGlyGlnAspAlaPheGlnGlnGlnGlnAla 84
  ::::::::::::::::::::
851 AGTATTCTCGATATCTCTTACAGTAACCTAAATCAGGAGATTAGTAACAT 802
  85 ValMetMetAspGlnLysAlaGlyLeuGlyGlnThr..... 97
  ::::::::::::::::::::
801 ACAGGATGATCAAAATTTAGCCCTTACAATAATTAATCTCAATGATCA 752
  98 .....TyrPro..AlaGlnGlyProProMetGlnGlyGlyP 109
  ::::::::::::::::::::
751 CTCGCTACCAAGATACCCCAATGCTGTAGATGCCATCAACAGAGTGTG 702
  109 heHisLeuGlnGlyLysIleSerProSerPheAsnSer...MetMetAsn 124
  ::::::::::::::::::::
701 AAGGACTATATGACACAGCAGCCATCCACCCAGTGGCTCACTTAACCAA 652
  125 MetAsnGlnGlnGlyAsnPheProLeuGln..... 134
  ::::::::::::::::::::
651 ATGAACACAGAAATATGATGATCTTACAGCTTCAAGGAACTTATGCTTC 602
  135 .....GlyMetHisProArgAlaAsnGlnMetArgProArgGlnAsnT 149
  ::::::::::::::::::::

```

```

601 TCCACCTCCCATGTCACCCATGAAGCAATGATATCCAGAGGCACTC 552
149 hProlGlnLeuArg 154
  ::::::::::::::::::::
551 CTCCTCCACAGTCAAG 535

```

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-870-693-8

seq_documentation_block:

```

: Sequence 8, Application US/08870693
: Patent No. 5866338
: GENERAL INFORMATION:
: APPLICANT: Hartwell, Leland H.
: APPLICANT: Weinert, Ted A.
: APPLICANT: Pilon, Sharon E.
: APPLICANT: Groudine, Mark T.
: TITLE OF INVENTION: Cell Cycle Checkpoint Genes
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
: STREET: 1420 Fifth Ave., Suite 2800
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98101-2347
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/870,693
: FILING DATE:
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/198,446
: FILING DATE: February 18, 1994
: APPLICATION NUMBER: PCT/US93/04458
: FILING DATE: May 12, 1993
: APPLICATION NUMBER: US 07/884,426
: FILING DATE: May 14, 1992
: APPLICATION NUMBER: US 07/882,051
: FILING DATE: May 12, 1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Shelness, Diana K.
: REGISTRATION NUMBER: 35,356
: REFERENCE/DOCKET NUMBER: FHCRI10798
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 206-682-8100
: TELEFAX: 206-224-0779
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1896 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: DESCRIPTION: human RAD9compa cDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
US-08-870-693-8

```

```

alignment_scores:
Quality: 83.00      Length: 189
Ratio: 0.988       Gaps: 8
Percent Similarity: 44.444      Percent Identity: 25.926

```

alignment_block:

US-09-041-994-2_COPY_1018_1179 x US-08-870-693-8/rev ..

Align seq 1/1 to reverse of: US-08-870-693-8 from: 1 to: 1896

```

3 ValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLeuAspAs 19
  :::::::::::::::::::: ||| ::| ::| ::| ::| ::| ::|
1074 TTGTCCTCCAGTGGCCCGAGAGTCCAGCATGGCCACCTCTGCGCTCA 1025
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
19 PleuVal.....GlyProSerAsnLeuGlnGlyGlns 31
  ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
1024 CTCGCTGCAGACGTTCCATCCACCCCTCTACTGCTCTCCATGAGAAAT 975
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
31 er.....AspGluArgAla 35
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
974 CCGTTGCCACAGTCCCATTTCTCCCGAATCCTCCCAAGAGGGGCT 925
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
36 LeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGlyLe 52
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
924 GTTAGGCCCGCAACCCCTTAACCTTAGTCTCGGAGC..... 889
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
52 uGluGluIleAspArgAlaLeuGlyIleProGluLeuValAsnGlnGly 69
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
888 .....CAGACAGTCCCTCTCTCTACTATAACACACTCAGGCGC 852
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
69 In..AlaLeuGluProLysGlnAspAlaPheGlnGlyGlnGluAlaAla 84
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
851 AGTATCTCGATATCTTACAGTAACTTAATCAGGATTAAGTAAACAAT 802
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
85 ValMetMetAspGlnLysAlaGlyLeuTyGlyGlnThr..... 97
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
801 ACAGGGATGAATCAAAATTTAGGCTTACAAATTAATCTCAATGAATCA 752
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
98 .....TyPro...AlaGlnGlyProPrometGlnGlyLys 109
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
751 GTCCGTACCAAGATACCCCAATGCTGTAGATTCCTCAACAGCTGCTC 702
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
109 heHisLeuGlnGlyGlnSerProSerPheAsnSer...MetMetAsnGln 124
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
701 AAGCACTAATGACACAGCAGCCATCCACCCAGTGGCTCACTTAACCAA 652
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
125 MetAsnGlnGlnGlyAsnPheProLeuGln..... 134
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
651 ATGAAACACACAAACTATGATCTTACAGCCTCAGGCACTTAATGCGCTC 602
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
135 .....GlyMetHisProArGlnAlaAsnIleMetArgProArgThrAsn 149
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
601 TTCACCTCCCATGTCACCATGAAAGCAATGAGTAATCCAGGAGCAGCTC 552
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
149 hrProLysGlnLeuArg 154
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
551 CTCCTCCACAAAGTCAGC 535
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|

```

seq_name: /cgn2_6/ptodata/1/lna/5B_COMB.seq:US-08-525-742-5

```

seq_documentation_block:
; Sequence 5, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Seki, Sakiko
; APPLICANT: Onawa, Ikuroh
; APPLICANT: Funato, Hiroh
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigem
; APPLICANT: Takahashi, Kiyochiro
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
; ADDRESSEE: NAUGHTON
; STREET: 1725 K Street, Suite 1000

```

```

CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,742
FILING DATE: 25-SEP-1995
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 05-074139
FILING DATE: 31-MAR-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 05-245625
FILING DATE: 30-SEP-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/JP94/00541
FILING DATE: 31-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mclelland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 950811
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-659-2930
TELEFAX: 202-8870357
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1945 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 97..1038
US-08-525-742-5

```

```

alignment_scores:
Quality: 79.00 Length: 154
Ratio: 1.113 Gaps: 6
Percent Similarity: 46.104 Percent Identity: 24.026

```

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-08-525-742-5 ..

Align seq 1/1 to: US-08-525-742-5 from: 1 to: 1945

```

21 ValGlyProProSerAsnLeuGlnGlyGlnSerAspGluArgAlaLeu 37
  ||| ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
367 GTCACCTCCAACTCAACAAGCTGAAGTTAAGACTGAACAATTAATGGCAC 416
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
37 uAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuGln 54
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
417 ACAATTAGTAACAACGTAGTGTAGTCAACATCAAGTCGAGTACTGAC 466
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
54 LuIleAsp..... 56
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
467 AAGTTGAAGGTATTATTACCTCTGTGTCACACACCAACGGAATCGCT 516
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
57 .....ArgAlaLeuGlyIleProGluLeuValAsnGlnGlyGln 70
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
517 CCACTCTCTTCCACCAATGGGTAGTCTAAGTTATTAGTCCAAACAGC 566
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
70 aLeuGluProLys.....GlnAspAlaPheGlnGlyGln 82
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
567 TGGTCATCCACACACGACGACGCTCCGATGAATGCTCAATCCAGGTCAC 616
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
82 LuIleAlaValMetMetAspGlnLysAlaGlyLeuTyGlyGlnThrTy 98
  ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|

```

```
617 CACGTCCTCCAAATGGGTAGTCCTAAG.....TTATTAGGTCCAAACCAA 660
      |||   ...   |||   |||   ...
99  ProAlaGlnGlyProPromeGlnGlyGly...PheHisLeuGlnGlyG1 114
      :::::   |||||   :::::   :::::   |||||
661 GCTGGTCATCCAAAGACCATGCCAAATGCTCCACAACCAACCAAGGTCC 710
      |||
114 nSerProSerPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlyAsnP 131
      |||
711 AAGACCA.....ATGAACCCCTCAAGGCAT. 735
      |||
131 heProLeuGlnGlyMetHisProArgAlaAsnIleMetArgProArgThr 147
      |||||   ...   |||||
736 .....CCTCGTCCTGGAACACGCTGGCCACGACCT 765
      |||
148 AsnThrProLys 151
      |||   |||   :::::
766 AACGGCCACAA 777
```



```
51 G1yleuGlulileaspargalaleuglyileProGluLeuValasn1 67
|||||
3287 GGCGTGAAGAAATTGACAGAGCTTTGGGATTCCTGAACCTGTCAATCA 3336
67 nglYGlAlAlaleuGluProLySGlnAspAlarPheGlnGlnAla1a 84
|||||
3337 GGGACAGGATTTAGAGCCAAACAGGATGCTTTCAGAGCCAGAAAGCCAG 3386
84 lavalmetetaspGlnLysAlaGlyLeuTyrglyInThrTyrrProAla 100
|||||
3387 CAGTAATGATGATCAGAAAGCCAGGATTTATATGGACAGACATACCAACA 3436
101 GlnGlyProProwetGlnGlyPheHnslsleuGlnGlyGlnserProse 117
|||||
3437 CAGGGCCCTCCATGACAGAGGCTTTCATCTTCAGAGCAATCACACATC 3486
117 rPheasnsermetetaspGlnmetasnGlnGlnGlnGlnGlnGlnGln 134
|||||
3487 TTTTAACCTATGATGAATCAGATGAACAGCAAGCAATTTTCCCTCTCC 3536
134 lnglymethisProArGAlaasnIleMetArProArGThrasnThrPro 150
|||||
3537 MAGAATGCACCCACAGCCACACATCATGAGACCCCGACAAACACCCCC 3586
151 LysGlnLeuArgMetGlnLeuGlnGlnGlnArgLeuGln 162
|||||
3587 AAGCACTTAGAATGCAGCTTCAGACAGGCTGCAG 3622

seq_name: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:US-09-513-066-1

seq_documentation_block:
: Sequence 1, Application US/09513066
: GENERAL INFORMATION:
: APPLICANT: Chen, J. Don
: APPLICANT: Leo, Christopher
: APPLICANT: Li, Hui
: TITLE OF INVENTION: NOVEL USES FOR THE RAC3 TRANSCRIPTIONAL COACTIVATOR OF
: FILE REFERENCE: UMG-026CP
: CURRENT APPLICATION NUMBER: US/09/513_066
: PRIOR FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: USSN 09/041,994
: PRIOR FILING DATE: 1998-03-13
: PRIOR APPLICATION NUMBER: USSN 60/073,674
: NUMBER OF SEQ ID NOS: 37
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 1
: LENGTH: 4496
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (86)..(4330)
US-09-513-066-1

Alignment_scores:
Quality: 845.00 Length: 162
Ratio: 5.216 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-09-513-066-1 ..
Align seg 1/1 to: US-09-513-066-1 from: 1 to: 4496
1 G1uGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSer1 17
|||||
3137 GAACAGATTCTCATGTCACACTCAAAATAGCCCTTCCTTAGGAATTCCT 3186
17 unspaspleuValGlyProProserAsnleuGlnGlnGlnserAspGluA 34
```

```
3187 GGATGATCTTTGGGCCACCTTCCAACTGGAAGGCCAGAGTACGAAA 3236
34 rglAlaleuAspGlnIleuHnslsThrleuSerAsnThrAspAlarP 50
|||||
3237 GAGCATTTATTTGGACAGCTGCACACTCTTCTCAGCAACAGATGGGACA 3286
51 G1yleuGlulileaspargalaleuglyileProGluLeuValasn1 67
|||||
3287 GGCGTGAAGAAATTGACAGAGCTTTGGGATTCCTGAACCTGTCAATCA 3336
67 nglYGlAlAlaleuGluProLySGlnAspAlarPheGlnGlnAla1a 84
|||||
3337 GGGACAGGATTTAGAGCCAAACAGGATGCTTTCAGAGCCAGAAAGCCAG 3386
84 lavalmetetaspGlnLysAlaGlyLeuTyrglyInThrTyrrProAla 100
|||||
3387 CAGTAATGATGATCAGAAAGCCAGGATTTATATGGACAGACATACCAACA 3436
101 GlnGlyProProwetGlnGlyPheHnslsleuGlnGlyGlnserProse 117
|||||
3437 CAGGGCCCTCCATGACAGAGGCTTTCATCTTCAGAGCAATCACACATC 3486
117 rPheasnsermetetaspGlnmetasnGlnGlnGlnGlnGlnGlnGln 134
|||||
3487 TTTTAACCTATGATGAATCAGATGAACAGCAAGCAATTTTCCCTCTCC 3536
134 lnglymethisProArGAlaasnIleMetArProArGThrasnThrPro 150
|||||
3537 MAGAATGCACCCACAGCCACACATCATGAGACCCCGACAAACACCCCC 3586
151 LysGlnLeuArgMetGlnLeuGlnGlnGlnArgLeuGln 162
|||||
3587 AAGCACTTAGAATGCAGCTTCAGACAGGCTGCAG 3622

seq_name: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:US-60-068-511-1

seq_documentation_block:
: Sequence 1, Application US/60068511
: GENERAL INFORMATION:
: APPLICANT: Suen, Chen-Shian
: APPLICANT: Frell, Donald E.
: APPLICANT: Lytle, Richard C.
: TITLE OF INVENTION: Cloning and Expression of a Nuclear
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESS: American Home Products Corporation
: STREET: One Campus Drive
: CITY: Parsippany
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07054
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/60/068,511
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Walsh, Andrea C.
: REGISTRATION NUMBER: 34,988
: REFERENCE/DOCKET NUMBER: 97243-00
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 973-683-2169
: TELEFAX: 973-683-4117
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4789 base pairs
: TYPE: nucleic acid
```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 185..4750
; US-60-068-511-1

```

```

alignment_scores:
  Quality: 845.00      Length: 162
  Ratio: 5.216        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

alignment_block:

```
US-09-041-994-2_COPY_1018_1179 x US-60-068-511-1 ..
```

```
Align seg 1/1 to: US-60-068-511-1 from: 1 to: 4789
```

```

1 GluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLe 17
  |||||||
3236 GAACAAGTTTCTCATGGCAGCTCAAAATAGGCTCTTCTTAGAATTCCT 3285
17 uAspAspLeuValGlyProProSerAsnLeuGluGlnSerAspGlnA 34
  |||||||
3286 GGATGATCTTGTGGGCCACCTTCCAACTGGAAGGCCAGAGTGAAGCAA 3335
34 rGAlaLeuLeuAspGlnLeuHisThrLeuSerAsnThrAspAlaThr 50
  |||||||
3336 GAGCATTTATTTGACAGCTGCACACTTCTTCACCAACACAGATGCCACA 3385
51 GlyLeuGluGlnIleAspArgAlaLeuGlyIleProGluLeuValAsnG 67
  |||||||
3386 GGCTTGAAGAAATTGACAGAGCTTTGGGCATTCCTGAACCTTGCAATCA 3435
67 nGlyGlnAlaLeuGluProLysGlnAspAlaPheGlnGlyGlnIuaAla 84
  |||||||
3436 GGGCAGGCAATTTAGAGCCCAAAACAGAGATGCTTCCAAAGGCCAAGAAC 3485
84 IValMetMetAspGlnLysAlaGlyLeuTyrglyGlnThrTyProAla 100
  |||||||
3486 CAGTAATGATGATCAGAAAGCAGGATTATATGACACACATACCCACCA 3535
101 GlnGlyProPrometGlnGlyPheHisLeuGlnGlyAsnProLeuG 117
  |||||||
3536 CAGGGGCTCCAAATGCAAGAGGCTTTTCATCTTCAGGACAAATACCATC 3585
117 rPheAsnSerMetLeuAsnGlnMetAsnGlnGlnIleAsnProLeuG 134
  |||||||
3586 TTTTAACCTCTATGATGATCAGATGACCAAGCAAGCAATTTCTCTCTCC 3635
134 InGlyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
  |||||||
3636 AAGGAATGACCCACGAGCCAAACATCATGAGACCCCGGACAAACACCCC 3685
151 LysGlnLeuArgMetGlnLeuGlnIleArgLeuGln 162
  |||||||
3686 AAGCAACTTAGAATGCAGCTTCACGACAGAGCTGCAG 3721

```

```
seq_name: /cgn2_6/ptodata/1/pna/US094_COMB.seq:US-09-440-612-1
```

seq_documentation_block:

```

; Sequence 1, Application US/09440612
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-3 EXPRESSION
; FILE REFERENCE: RTS-0042
; CURRENT APPLICATION NUMBER: US/09/440,612

```

```

; CURRENT FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 6754
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (184)..(4422)
; US-09-440-612-1

```

```

alignment_scores:
  Quality: 845.00      Length: 162
  Ratio: 5.216        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

alignment_block:

```
US-09-041-994-2_COPY_1018_1179 x US-09-440-612-1 ..
```

```
Align seg 1/1 to: US-09-440-612-1 from: 1 to: 6754
```

```

1 GluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLe 17
  |||||||
3220 GAACAAGTTTCTCATGGCAGCTCAAAATAGGCTCTTCTTAGAATTCCT 3269
17 uAspAspLeuValGlyProProSerAsnLeuGluGlnSerAspGlnA 34
  |||||||
3270 GGATGATCTTGTGGGCCACCTTCCAACTGGAAGGCCAGAGTGAAGCAA 3319
34 rGAlaLeuLeuAspGlnLeuHisThrLeuSerAsnThrAspAlaThr 50
  |||||||
3320 GAGCATTTATTTGACAGCTGCACACTTCTTCACCAACACAGATGCCACA 3369
51 GlyLeuGluGlnIleAspArgAlaLeuGlyIleProGluLeuValAsnG 67
  |||||||
3370 GGCTTGAAGAAATTGACAGAGCTTTGGGCATTCCTGAACCTTGCAATCA 3419
67 nGlyGlnAlaLeuGluProLysGlnAspAlaPheGlnGlyGlnIuaAla 84
  |||||||
3420 GGGCAGGCAATTTAGAGCCCAAAACAGAGATGCTTCCAAAGGCCAAGAAC 3469
84 IValMetMetAspGlnLysAlaGlyLeuTyrglyGlnThrTyProAla 100
  |||||||
3470 CAGTAATGATGATCAGAAAGCAGGATTATATGACACACATACCCACCA 3519
101 GlnGlyProPrometGlnGlyPheHisLeuGlnGlyAsnProLeuG 117
  |||||||
3520 CAGGGGCTCCAAATGCAAGAGGCTTTTCATCTTCAGGACAAATACCATC 3569
117 rPheAsnSerMetLeuAsnGlnMetAsnGlnGlnIleAsnProLeuG 134
  |||||||
3570 TTTTAACCTCTATGATGATCAGATGACCAAGCAAGCAATTTCTCTCTCC 3619
134 InGlyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
  |||||||
3620 AAGGAATGACCCACGAGCCAAACATCATGAGACCCCGGACAAACACCCC 3669
151 LysGlnLeuArgMetGlnLeuGlnIleArgLeuGln 162
  |||||||
3670 AAGCAACTTAGAATGCAGCTTCACGACAGAGCTGCAG 3705

```

```
seq_name: /cgn2_6/ptodata/1/pna/US094_COMB.seq:US-09-440-612-3
```

seq_documentation_block:

```

; Sequence 3, Application US/09440612A
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-3 EXPRESSION
; FILE REFERENCE: RTS-0042
; CURRENT APPLICATION NUMBER: US/09/440,612A
; CURRENT FILING DATE: 1999-11-15

```

```
;; NUMBER OF SEQ ID NOS: 49
;; SEQ ID NO 3
;; LENGTH: 6754
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (184)..(4422)
US-09-440-612-3
```

alignment_scores:

```
Quality: 845.00 Length: 162
Ratio: 5.216 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

```
US-09-041-994-2_COPY_1018_1179 x US-09-440-612-3 ..
```

```
Align seg 1/1 to: US-09-440-612-3 from: 1 to: 6754
```

```
1 GUGLNVAlSerHISglYThrGlnAsnArgProLeuLeuArgAsnSerle 17
|||||
3220 GAACAAGTTTCTCATGCGACTCAAAATAGGCGCTTCTTAGGAATTCCT 3269
17 uaspaSpleuValGlyProProSerAsnLeuGluGlnSerAspGluA 34
|||||
3270 GCATGATCTTGTGGGCCACCTTCCAACTGGAAAGCCAGAGTGAAGAA 3319
34 rglAlaLeuLeuAspGlnLeuHISThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3320 GAGCATTTATGGACCGACGTCGACACCTTCTCAGCAACACAGATGCCACA 3369
51 GYLLeuGluGluLeuAspArgAlaLeuGlyIleProGluLeuValAsnG1 67
|||||
3370 GCGCTGGAAAGAAATTCACAGAGCTTGGGCATTCCTGAAACTGTCAATCA 3419
67 nGlyGlnAlaLeuGluProLySGlnAspAlaPheGlnGlyGlnGluAlaA 84
|||||
3420 GGGACAGGCAATTAGACCCCAACAGGATGCTTCCAAAGGCCAAAGAGCAG 3469
84 lavalMetMetAspGlnLysAlaGlyLeuTyrglyGlnThrTyrrProAla 100
|||||
3470 CAGTAAATGATGATCAGAAAGCGAGATTAATGACAGACATACCCAGCA 3519
101 GlnGlyProPrometGlnGlyIlePheHISLeuGlnGlyGlnSerProse 117
|||||
3520 CAGGGGCTTCATATGCAAGAGGCTTTCATCTTCAGGACCAATCACCATC 3569
117 rPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlnLysAsnPheProLeuG 134
|||||
3570 TTTTAACTTATGATGAAATCAGATGAACCAAGCAATTTTCCCTCC 3619
134 InGlyMetHISProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
|||||
3620 AAGGAATGCAACCCAGACGACCAATCATGAGACCCCGGACAAACACCCCC 3669
151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
|||||
3670 AAGCAACTTAGAATGACGCTTCAGCAGAGGCTGCAG 3705
```

```
seq_name: /cgn2_6/ptodata/1/pna/US091_COMB.seq:US-09-125-635-1
```

seq_documentation_block:

```
;; Sequence 1, Application US/09125635
;; GENERAL INFORMATION:
;; APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE
;; TITLE OF INVENTION: AIB1, A novel steroid receptor co-activator
;; FILE REFERENCE: 49944
;; CURRENT APPLICATION NUMBER: US/09/125, 635
;; CURRENT FILING DATE: 1998-08-21
;; PRIOR APPLICATION NUMBER: 60/049,728
;; PRIOR FILING DATE: 1997-06-17
```

```
;; NUMBER OF SEQ ID NOS: 12
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 1
;; LENGTH: 6835
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (201)..(4463)
US-09-125-635-1
```

alignment_scores:

```
Quality: 845.00 Length: 162
Ratio: 5.216 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

```
US-09-041-994-2_COPY_1018_1179 x US-09-125-635-1 ..
```

```
Align seg 1/1 to: US-09-125-635-1 from: 1 to: 6835
```

```
1 GUGLNVAlSerHISglYThrGlnAsnArgProLeuLeuArgAsnSerle 17
|||||
3252 GAACAAGTTTCTCATGCGACTCAAAATAGGCGCTTCTTAGGAATTCCT 3301
17 uaspaSpleuValGlyProProSerAsnLeuGluGlnSerAspGluA 34
|||||
3302 GCATGATCTTGTGGGCCACCTTCCAACTGGAAAGCCAGAGTGAAGAA 3351
34 rglAlaLeuLeuAspGlnLeuHISThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3352 GAGCATTTATGGACCGACGTCGACACCTTCTCAGCAACACAGATGCCACA 3401
51 GYLLeuGluGluLeuAspArgAlaLeuGlyIleProGluLeuValAsnG1 67
|||||
3402 GCGCTGGAAAGAAATTCACAGAGCTTGGGCATTCCTGAAACTGTCAATCA 3451
67 nGlyGlnAlaLeuGluProLySGlnAspAlaPheGlnGlyGlnGluAlaA 84
|||||
3452 GGGACAGGCAATTAGACCCCAACAGGATGCTTCCAAAGGCCAAAGAGCAG 3501
84 lavalMetMetAspGlnLysAlaGlyLeuTyrglyGlnThrTyrrProAla 100
|||||
3502 CAGTAAATGATGATCAGAAAGCGAGATTAATGACAGACATACCCAGCA 3551
101 GlnGlyProPrometGlnGlyIlePheHISLeuGlnGlyGlnSerProse 117
|||||
3552 CAGGGGCTTCATATGCAAGAGGCTTTCATCTTCAGGACCAATCACCATC 3601
117 rPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlnLysAsnPheProLeuG 134
|||||
3602 TTTTAACTTATGATGAAATCAGATGAACCAAGCAATTTTCCCTCC 3651
134 InGlyMetHISProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
|||||
3652 AAGGAATGCAACCCAGACGACCAATCATGAGACCCCGGACAAACACCCCC 3701
151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
|||||
3702 AAGCAACTTAGAATGACGCTTCAGCAGAGGCTGCAG 3737
```

```
seq_name: /cgn2_6/ptodata/1/pna/US093_COMB.seq:US-09-359-922-3962
```

seq_documentation_block:

```
;; Sequence 3962, Application US/09359922
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc.
;; TITLE OF INVENTION: NOVEL CONFIGS OBTAINED FROM VARIOUS CDNA
;; FILE REFERENCE: 20411-752CON1
;; CURRENT APPLICATION NUMBER: US/09/359, 922
;; CURRENT FILING DATE: 1998-07-22
```


EARLIER APPLICATION NUMBER: US 09/205,155
EARLIER FILING DATE: 1998-12-03
NUMBER OF SEQ ID NOS: 13203
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 3962
LENGTH: 6855
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(6855)
OTHER INFORMATION: n = A,T,C or G
US-09-359-922-3962

alignment_scores:
Quality: 845.00 Length: 162
Ratio: 5.216 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-041-994-2_COPY_1018_1179 x US-09-359-922-3962 ..

Align seg 1/1 to: US-09-359-922-3962 from: 1 to: 6855

```
1  GluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLe 17
|||||
3302  GAGCATGATCTTTGGGCCACCTTCCAACTGGAAGCCAGAGTGCAGAA 3351
|||||
17  uAspAspLeuValGlyProProSerAsnLeuGlnGlyInserAspGluA 34
|||||
3302  GAGCATGATCTTTGGGCCACCTTCCAACTGGAAGCCAGAGTGCAGAA 3351
|||||
34  rGAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3352  GAGCATGATCTTTGGGCCACCTTCCAACTGGAAGCCAGAGTGCAGAA 3401
|||||
51  GlyLeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnG 67
|||||
3402  GGCCTGGAGAAATTCAGACAGCTTTGGGCATCTCTGAACCTTGCAATCA 3451
|||||
67  nGlyGlnAlaLeuGlnProLysGlnAspAlaPheGlnGlyGlnIleAla 84
|||||
3452  GGGACAGGCAATTAGAGCCCAAGAGATGCTTCCAAAGCCAGAGAGCAG 3501
|||||
84  lAlaMetMetAspGlnLysAlaGlyLeuTyrglyGlnThrTyProAla 100
|||||
3502  CAGTAATGATGATCAGAAAGCAGGATTAATGACAGACATATCCACACA 3551
|||||
101  GlnGlyProPrometGlnGlyPheHisLeuGlnGlyInserProSe 117
|||||
3552  CAGGGGCTCCCAATGCAAGAGGCTTTCATCTTCAGGCAATCAATCCATC 3601
|||||
117  rPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlnLysAsnPheProLeu 134
|||||
3602  TTTTAACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3651
|||||
134  lnglyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
|||||
3652  AAGGAATGACACCCAGAGCCCAACATCATGAGACCCGAGCAAAACACCCC 3701
|||||
151  LysGlnLeuArgMetGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 162
|||||
3702  AAGCACTTAGAATGACCTTCAGCAGAGGCTGCAG 3737
```

seq_name: /cgn2_6/ptodata/1/pna/US093_COMB.seq:us-09-359-922-3962

seq_documentation_block:

; Sequence 3962, Application US/09359922A
; GENERAL INFORMATION:
; APPLICANT: Leshkowitz, Dena
; TITLE OF INVENTION: NOVEL CONFIGS OBTAINED FROM VARIOUS CDNA

TITLE OF INVENTION: LIBRARIES
FILE REFERENCE: 20411-752CON1
CURRENT APPLICATION NUMBER: US/09/359, 922A
CURRENT FILING DATE: 1999-07-22
EARLIER APPLICATION NUMBER: US 09/205,155
EARLIER FILING DATE: 1998-12-03
EARLIER APPLICATION NUMBER: US 09/034,341
EARLIER FILING DATE: 1998-02-13
NUMBER OF SEQ ID NOS: 13203
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 3962
LENGTH: 6855
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(6855)
OTHER INFORMATION: n = A,T,C or G
US-09-359-922-3962

alignment_scores:
Quality: 845.00 Length: 162
Ratio: 5.216 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-041-994-2_COPY_1018_1179 x US-09-359-922-3962 ..

Align seg 1/1 to: US-09-359-922-3962 from: 1 to: 6855

```
1  GluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLe 17
|||||
3302  GAGCATGATCTTTGGGCCACCTTCCAACTGGAAGCCAGAGTGCAGAA 3351
|||||
17  uAspAspLeuValGlyProProSerAsnLeuGlnGlyInserAspGluA 34
|||||
3302  GAGCATGATCTTTGGGCCACCTTCCAACTGGAAGCCAGAGTGCAGAA 3351
|||||
34  rGAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3352  GAGCATGATCTTTGGGCCACCTTCCAACTGGAAGCCAGAGTGCAGAA 3401
|||||
51  GlyLeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnG 67
|||||
3402  GGCCTGGAGAAATTCAGACAGCTTTGGGCATCTCTGAACCTTGCAATCA 3451
|||||
67  nGlyGlnAlaLeuGlnProLysGlnAspAlaPheGlnGlyGlnIleAla 84
|||||
3452  GGGACAGGCAATTAGAGCCCAAGAGATGCTTCCAAAGCCAGAGAGCAG 3501
|||||
84  lAlaMetMetAspGlnLysAlaGlyLeuTyrglyGlnThrTyProAla 100
|||||
3502  CAGTAATGATGATCAGAAAGCAGGATTAATGACAGACATATCCACACA 3551
|||||
101  GlnGlyProPrometGlnGlyPheHisLeuGlnGlyInserProSe 117
|||||
3552  CAGGGGCTCCCAATGCAAGAGGCTTTCATCTTCAGGCAATCAATCCATC 3601
|||||
117  rPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlnLysAsnPheProLeu 134
|||||
3602  TTTTAACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3651
|||||
134  lnglyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
|||||
3652  AAGGAATGACACCCAGAGCCCAACATCATGAGACCCGAGCAAAACACCCC 3701
|||||
151  LysGlnLeuArgMetGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 162
|||||
3702  AAGCACTTAGAATGACCTTCAGCAGAGGCTGCAG 3737
```

seq_name: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:us-60-209-009-198

```
seq_documentation_block:
: Sequence 198, Application US/60209009
: GENERAL INFORMATION:
: APPLICANT: Paris, Mary
: APPLICANT: Pearson, Cecelia I.
: TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
: FILE REFERENCE: PA-0027 P
: CURRENT APPLICATION NUMBER: US/60/209, 009
: NUMBER OF SEQ ID NOS: 501
: SOFTWARE: PERL Program
: SEQ ID NO 198
: LENGTH: 7116
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc-feature
: OTHER INFORMATION: Incyte ID No: 1094199.1
: NAME/KEY: unsure
: LOCATION: 3941-3993, 5899-5939, 6951
: OTHER INFORMATION: a, t, c, g, or other
: US-60-209-009-198
```

alignment_scores:

Quality:	845.00	Length:	162
Ratio:	5.216	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-041-994-2_COPY_1018_1179 x US-60-209-009-198 ..

Align seg 1/1 to: US-60-209-009-198 from: 1 to: 7116

```
1 1GUG1NValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSer17
|||||
3252 GAACAGATTTCATGCGACTCAAAATAGGCGCTCTTAGGAATCCCT 3301
17 uAspAspLeuValGlyProProSerAsnLeuGlnGlyGlnSerAspGlu 34
|||||
3302 GGATGATCTTGTGGGCCACCTTCCAACTGGAAGGCCAGAGTGCAGAA 3351
34 rGAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3352 GACCATTAATGGACGACGCTGCACACTCTTCAGCAACAGATGCCACA 3401
51 G1yLeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnG1 67
|||||
3402 GGCCTGGAAGAAATTGACAGAGCTTTGGGCATTCTGAACTTGTCAATCA 3451
67 nG1yGlnAlaLeuGluProLySGlnAspAlaPheGlnGlnGluAla 84
|||||
3452 GGGACAGGCAATTAGAGCCCAACAGAGATGCTTCCAAAGCCCAAGAGCAG 3501
84 lAvalMetLeuAspGlnLySAlaGlyLeuTyGlyGlnThrTyProAla 100
|||||
3502 CAGTAATGATGGATCAGAAAGCGAGATTATATGACAGACATACCCAGCA 3551
101 GlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlyGlnSerProse 117
|||||
3552 CAGGGGCTCCCAATGCAAGAGAGGCTTTCATCTTCAGGACAAATCACACATC 3601
117 rPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlyAsnPheProLeuG 134
|||||
3602 TTTTAACCTATGATGAATCAGATGAACCAAGCAATTTTCCCTCC 3651
134 lnglyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
|||||
3652 AAGGAATGACACCCAGACGCAACATCATAGAGCCCGGACAAACACACCCC 3701
151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
|||||
3702 AAGCAACTTAGAATGCAAGCTTCAGCAGAGGCTGCAG 3737
```

seq_name: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:US-60-213-360-7991

```
seq_documentation_block:
: Sequence 7991, Application US/60213360
: GENERAL INFORMATION:
: APPLICANT: Morris, MacDonald
: APPLICANT: Lal, Preethi
: APPLICANT: Diep, Dinh
: TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
: TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Poly
: TITLE OF INVENTION: Identified Thereby
: FILE REFERENCE: GX-0014 P
: CURRENT APPLICATION NUMBER: US/60/213, 360
: CURRENT FILING DATE: 2000-06-21
: NUMBER OF SEQ ID NOS: 8347
: SOFTWARE: PERL Program
: SEQ ID NO 7991
: LENGTH: 7116
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc-feature
: OTHER INFORMATION: Incyte ID No: 1094199.1
: NAME/KEY: unsure
: LOCATION: 3941-3993, 5899-5939, 6951
: OTHER INFORMATION: a, t, c, g, or other
: US-60-213-360-7991
```

alignment_scores:

Quality:	845.00	Length:	162
Ratio:	5.216	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-041-994-2_COPY_1018_1179 x US-60-213-360-7991 ..

Align seg 1/1 to: US-60-213-360-7991 from: 1 to: 7116

```
1 1GUG1NValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSer17
|||||
3252 GAACAGATTTCATGCGACTCAAAATAGGCGCTCTTAGGAATCCCT 3301
17 uAspAspLeuValGlyProProSerAsnLeuGlnGlyGlnSerAspGlu 34
|||||
3302 GGATGATCTTGTGGGCCACCTTCCAACTGGAAGGCCAGAGTGCAGAA 3351
34 rGAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3352 GACCATTAATGGACGACGCTGCACACTCTTCAGCAACAGATGCCACA 3401
51 G1yLeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAsnG1 67
|||||
3402 GGCCTGGAAGAAATTGACAGAGCTTTGGGCATTCTGAACTTGTCAATCA 3451
67 nG1yGlnAlaLeuGluProLySGlnAspAlaPheGlnGlnGluAla 84
|||||
3452 GGGACAGGCAATTAGAGCCCAACAGAGATGCTTCCAAAGCCCAAGAGCAG 3501
84 lAvalMetLeuAspGlnLySAlaGlyLeuTyGlyGlnThrTyProAla 100
|||||
3502 CAGTAATGATGGATCAGAAAGCGAGATTATATGACAGACATACCCAGCA 3551
101 GlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlyGlnSerProse 117
|||||
3552 CAGGGGCTCCCAATGCAAGAGAGGCTTTCATCTTCAGGACAAATCACACATC 3601
117 rPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlyAsnPheProLeuG 134
|||||
3602 TTTTAACCTATGATGAATCAGATGAACCAAGCAATTTTCCCTCC 3651
134 lnglyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
```

```
|||||
3652 AAGGAATGCACCCAGAGCCACATCATGAGACCCCGACAAACACCCCC 3701
151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
|||||
3702 AAGCAACTTAGAATGACAGCTTCACAGAGAGGCTGCAG 3737

seq_name: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:US-09-652-123-8947

seq_documentation_block:
; Sequence 8947, Application US/09652123
; GENERAL INFORMATION:
; APPLICANT: Holzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1186-001
; CURRENT APPLICATION NUMBER: US/09/652,123
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,135
; NUMBER OF SEQ ID NOS: 9796
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8947
; LENGTH: 7888
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-123-8947
```

```
alignment_scores:
  Quality: 845.00      Length: 162
  Ratio: 5.216        Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-09-652-123-8947 ..

Align seg 1/1 to: US-09-652-123-8947 from: 1 to: 7888

```
1 GluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLe 17
|||||
3218 GACACAGTCTTCATGCGACACCAAAATGAGCCCTCTCTTGAAGATTCCT 3267
17 uAspAspLeuValGlyProProSerAsnLeuGlyGlnSerAspGlu 34
|||||
3268 GGATGATCTTGTGGCCACCTTCCAACTGGAAGGCCAGAGTACGAAA 3317
34 rGAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3318 GAGCATTAATGGACACAGCTGACACTCTTCTCAGACCAACAGATGCCACA 3367
51 GlyLeuGlnGluIleAspArgAlaLeuGlyTleProGluLeuValAsnG1 67
|||||
3368 GGCCTGGAGAAATTTGACAGAGCTTTGGCATTCCTGAACCTTGTCAATCA 3417
67 nGlyGlnAlaLeuGlnIleProLysGlnAspAlaPheGlnGlyGlnGlnAla 84
|||||
3418 GGGACAGGCAATTAGAGCCCAACAGAGATGCTTCCAAAGGCCAAAGACGAG 3467
84 lAValMetLeuAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAla 100
|||||
3468 CAGTAATGATGATGATCAGAGCAGAGATTAATGACAGACATACCAGCA 3517
101 GlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlnGlnSerProSe 117
|||||
3518 CAGGGGCTCCCAATGCAAGAGGCTTTCATCTTCAGGAGACATCACCATC 3567
117 rPheAsnSerMetLeuAsnGlnMetAsnGlnGlnGlyAsnPheProLeuG 134
|||||
3568 TTTTAACCTATGATGATCATCATGATGAACACGACCAATTTTCTCTCC 3617
134 lnglyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
|||||
```

```
3618 AAGGAATGCACCCAGAGCCACATCATGAGACCCCGACAAACACCCCC 3667
151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
|||||
3668 AAGCAACTTAGAATGACAGCTTCACAGAGAGGCTGCAG 3703

seq_name: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:US-09-652-127-9831

seq_documentation_block:
; Sequence 9831, Application US/09652127
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1183-001
; CURRENT APPLICATION NUMBER: US/09/652,127
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,134
; NUMBER OF SEQ ID NOS: 10475
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9831
; LENGTH: 7888
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-127-9831
```

```
alignment_scores:
  Quality: 845.00      Length: 162
  Ratio: 5.216        Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-09-652-127-9831 ..

Align seg 1/1 to: US-09-652-127-9831 from: 1 to: 7888

```
1 GluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLe 17
|||||
3218 GACACAGTCTTCATGCGACACCAAAATGAGCCCTCTCTTGAAGATTCCT 3267
17 uAspAspLeuValGlyProProSerAsnLeuGlyGlnSerAspGlu 34
|||||
3268 GGATGATCTTGTGGCCACCTTCCAACTGGAAGGCCAGAGTACGAAA 3317
34 rGAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3318 GAGCATTAATGGACACAGCTGACACTCTTCTCAGACCAACAGATGCCACA 3367
51 GlyLeuGlnGluIleAspArgAlaLeuGlyTleProGluLeuValAsnG1 67
|||||
3368 GGCCTGGAGAAATTTGACAGAGCTTTGGCATTCCTGAACCTTGTCAATCA 3417
67 nGlyGlnAlaLeuGlnIleProLysGlnAspAlaPheGlnGlyGlnGlnAla 84
|||||
3418 GGGACAGGCAATTAGAGCCCAACAGAGATGCTTCCAAAGGCCAAAGACGAG 3467
84 lAValMetLeuAspGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAla 100
|||||
3468 CAGTAATGATGATGATCAGAGCAGAGATTAATGACAGACATACCAGCA 3517
101 GlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlnGlnSerProSe 117
|||||
3518 CAGGGGCTCCCAATGCAAGAGGCTTTCATCTTCAGGAGACATCACCATC 3567
117 rPheAsnSerMetLeuAsnGlnMetAsnGlnGlnGlyAsnPheProLeuG 134
|||||
3568 TTTTAACCTATGATGATCATCATGATGAACACGACCAATTTTCTCTCC 3617
134 lnglyMetHisProArgAlaAsnIleMetArgProArgThrAsnThrPro 150
|||||
3618 AAGGAATGCACCCAGAGCCCAACATCATGAGACCCCGACAAACACCCCC 3667
```

151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
|||||
3668 AAGCAACTTAGATGAGCTTCAGCAGAGGCTGCAG 3703

seq_name: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:US-09-652-355-11058

seq_documentation_block:

; Sequence 11058, Application US/09652355
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1192-001
; CURRENT APPLICATION NUMBER: US/09/652,355
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,136
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 11227
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11058
; LENGTH: 7888
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-355-11058

alignment_scores:
Quality: 845.00 Length: 162
Ratio: 5.216 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-09-652-355-11058 ..

Align seg 1/1 to: US-09-652-355-11058 from: 1 to: 7888

1 GluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSer 17
|||||
3218 GAACAAGTTTCTCATGGCACTCAMAATAGGCTCTTAGGAATTCCT 3267
17 uaspaRpleuValGlyProProSerAsnLeuGlnGlnSerAspGlu 34
|||||
3268 GGAATGATCTTTGGGCCACCTTCCAACTGGAGGCCAGAGTGAGAAA 3317
34 rgaLaleuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3318 GAGCATTATTGGACAGCTGCACACTCTTCTCAGCAACAGATGCCACA 3367
51 GlyLeuGlnGluIleAspArgAlaLeuGlyIleProGlnLeuValAsnG 67
|||||
3368 GGCTTGGAGAAATTCACAGAGCTTTGGCATTCCTGAACTGTCAATCA 3417
67 nGlyGlnAlaLeuGluProLysGlnAspAlaPheGlnGlnGlnAla 84
|||||
3418 GGGACAGCGATTAGAGCCCAACAGATGCTTTCCAAAGGCCAAGAGCAG 3467
84 IavAlaMetCaspGlnLysAlaGlyLeuTyrglyGlnThrTyrrProAla 100
|||||
3468 CAGTAAATGATGATCAGAGCGAGCATTTATGACAGACATACCCAGCA 3517
101 GlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlnGlnSerPro 117
|||||
3518 CAGGGGCTTCCAATGCAGAGAGGCTTTTCATCTTCAGAGCAATCACCATC 3567
117 rPheAsnSerMetCaspGlnLysAlaGlnGlnGlnGlnGlnGlnGln 134
|||||
3568 TTTTAACTGATGATGAATCAGATGAACAGCAGCAAGCAATTTTCTCTCC 3617
134 lnglyMetHisProArgAlaAsnIleMetCargProArgThrAsnThrPro 150
|||||
3618 AAGGAATGACACCCAGCAGCAACATCATGAGACCCCGGACAAACACCCCC 3667

151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
|||||
3668 AAGCAACTTAGATGAGCTTCAGCAGAGGCTGCAG 3703

seq_name: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:US-09-663-693-1028

seq_documentation_block:

; Sequence 1028, Application US/09663693
; GENERAL INFORMATION:
; APPLICANT: Geating, David P.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1195-001
; CURRENT APPLICATION NUMBER: US/09/663,693
; CURRENT FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/154,986
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 1340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1028
; LENGTH: 7888
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-663-693-1028

alignment_scores:
Quality: 845.00 Length: 162
Ratio: 5.216 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-09-663-693-1028 ..

Align seg 1/1 to: US-09-663-693-1028 from: 1 to: 7888

1 GluGlnValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSer 17
|||||
3218 GAACAAGTTTCTCATGGCACTCAMAATAGGCTCTTAGGAATTCCT 3267
17 uaspaRpleuValGlyProProSerAsnLeuGlnGlnSerAspGlu 34
|||||
3268 GGAATGATCTTTGGGCCACCTTCCAACTGGAGGCCAGAGTGAGAAA 3317
34 rgaLaleuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3318 GAGCATTATTGGACAGCTGCACACTCTTCTCAGCAACAGATGCCACA 3367
51 GlyLeuGlnGluIleAspArgAlaLeuGlyIleProGlnLeuValAsnG 67
|||||
3368 GGCTTGGAGAAATTCACAGAGCTTTGGCATTCCTGAACTGTCAATCA 3417
67 nGlyGlnAlaLeuGluProLysGlnAspAlaPheGlnGlnGlnAla 84
|||||
3418 GGGACAGCGATTAGAGCCCAACAGATGCTTTCCAAAGGCCAAGAGCAG 3467
84 IavAlaMetCaspGlnLysAlaGlyLeuTyrglyGlnThrTyrrProAla 100
|||||
3468 CAGTAAATGATGATCAGAGCGAGCATTTATGACAGACATACCCAGCA 3517
101 GlnGlyProProMetGlnGlyGlyPheHisLeuGlnGlnGlnSerPro 117
|||||
3518 CAGGGGCTTCCAATGCAGAGAGGCTTTTCATCTTCAGAGCAATCACCATC 3567
117 rPheAsnSerMetCaspGlnLysAlaGlnGlnGlnGlnGlnGlnGln 134
|||||
3568 TTTTAACTGATGATGAATCAGATGAACAGCAGCAAGCAATTTTCTCTCC 3617
134 lnglyMetHisProArgAlaAsnIleMetCargProArgThrAsnThrPro 150
|||||
3618 AAGGAATGACACCCAGCAGCAACATCATGAGACCCCGGACAAACACCCCC 3667

|||||
3668 AAGCACTTAGAATGCAGCTTCAGCAGAGGCTGCAG 3703

seq_name: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:US-09-698-014-5135

seq_documentation_block:

: Sequence 5135, Application US/09698014
: GENERAL INFORMATION:
: APPLICANT: Gearing, David P.
: APPLICANT: Holtzman, Douglas A.
: APPLICANT: Kingsbury, Gillian A.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
: TITLE OF INVENTION: THEREFOR
: FILE REFERENCE: 1600, 2014-001
: CURRENT APPLICATION NUMBER: US/09/698,014
: CURRENT FILING DATE: 2000-10-27
: PRIOR APPLICATION NUMBER: 60/162,363
: PRIOR FILING DATE: 1999-10-29
: NUMBER OF SEQ ID NOS: 6098
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5135
: LENGTH: 7888
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-698-014-5135

alignment_scores:

Quality: 845.00 Length: 162
Ratio: 5.216 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-041-994-2_COPY_1018_1179 x US-09-698-014-5135

Align seg 1/1 to: US-09-698-014-5135 from: 1 to: 7888

1 GUGInValSerHisGlyThrGlnAsnArgProLeuLeuArgAsnSerLe 17
|||||
3218 GAACAGATTTCATGCGACGCTCAAAATAGGCCCTTCTTAGGAATTCCT 3267
17 uAspAspLeuValGlyProProSerAsnLeuGlnGlnSerAspGluA 34
|||||
3268 GGATGATCTTGTGGGCCACCTTCACACTGGAGGCCAGAGGACGAAA 3317
34 tGAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThr 50
|||||
3318 GAGCATTATTGGACGACGCTGCACACTCTTCAGCAACACAGATGCCACA 3367
51 G1yLeuG1uG1u1leAspArgAlaLeuG1y1leProG1uLeuValAsnG1 67
|||||
3368 GGCCTGGAAAGAAATGCAGAGCTTTGGGCATTCTGAACTTGTCATCA 3417
67 nG1yG1nAlaLeuGluProLySGlnAspAlaPheGlnGlnGlnAlaA 84
|||||
3418 GGGACAGGATAGAGCCCAACAGATGCTTCCAGGCCAAGACAGACAG 3467
84 lAvalMetMetAspGlnLySAlaGlyLeuTyrcLyGlnThrTyProAla 100
|||||
3468 CAGTAATGATGATGCAGAGCAGATATATGACAGACATACCCAGCA 3517
101 G1nG1yProProMetGlnG1yG1yPheHisLeuGlnG1yGlnSerProSe 117
|||||
3518 CAGGGGCTTCCAATGCAGAGAGGCTTTCATCTTCAGGACAATCACCATC 3567
117 rPheAsnSerMetMetAsnGlnMetAsnGlnGlnGlnLyAsnPheProLeuG 134
|||||
3568 TTTTAACCTATGATGAATCAGATGAACGACGACGATTTTCTCTGCC 3617
134 lngLyMetHisProArgAlaAsn1leMetArgProArgThrAsnThrPro 150
|||||
3618 AAGGAATGCACCCAGCAGCCACATCATGAGACCCCGACAAACACACCCC 3667

151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
|||||
3668 AAGCACTTAGAATGCAGCTTCAGCAGAGGCTGCAG 3703

151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
|||||
3686 AAGCACTTAGAATGACAGCTTCACAGCAGGCTGCAG 3721

seq_name: /cgn2_6/prodata/2/pna/us09_NEW_COMB.seq:us-09-540-212A-48616

seq_documentation_block:

; Sequence 48616, Application US/09540212A
; GENERAL INFORMATION:
; APPLICANT: Sellhauer, Jeffrey J.
; APPLICANT: Deleage, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Scuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF AIRWAY AND LUNG SYSTEM TISSUE
; FILE REFERENCE: PD-1034 CIP
; CURRENT APPLICATION NUMBER: US/09/540, 212A
; CURRENT FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 67551
; SOFTWARE: PERL Program
; SEQ ID NO 48616
; LENGTH: 274
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu01311169
US-09-540-212A-48616

alignment_scores:

Quality:	123.50	Length:	81
Ratio:	2.245	Gaps:	3
Percent Similarity:	67.901	Percent Identity:	40.741

alignment_block:

us-09-041-994-2_COPY_1018_1179 x us-09-540-212A-48616 ..

Align seg 1/1 to: us-09-540-212A-48616 from: 1 to: 274

35 AAlaLeuLeuArgGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrG1 51
|||||
27 GCTCTTCTTGACAGCTGGTATCCTCTTAGTGGAAGTTGAAAGTGA 76
51 yLeuGlnGluLeuAspArgAlaLeuGlyIleProGluLeuValAsnGln 68
||| |||
77 GCTAGCTGAAGTACAGCAGCTGCGAATTCACAAACTTGT...CAGG 123
68 yGlnAlaLeuGlnProLysGlnAspAlaPheGlnGlnGlnAlaAla 84
||| |||
124 GGGGTGTAGATGTATTATCAGAGATTTCCACACCAACAAGCAAGC 173
85ValMetMetAspGlnLysAlaGlyLeuTyrglyGlnThrTyrr 99
.....
174 CCACCTTTGATCATGAGAGAAAGACCCACCTTATTCACGCTTACTC 223

99 oAlaGlnGlyPro.....PrometGlnGly 107
::: |||
224 TTCCTCTTCTCTACTGCCATCTCCCTAGCCCTTTCCAAAGGC 266

seq_name: /cgn2_6/prodata/2/pna/us60_NEW_COMB.seq:us-60-248-505-387

seq_documentation_block:

; Sequence 387, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: c1000918
; CURRENT APPLICATION NUMBER: US/60/248, 505

; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 387
; LENGTH: 5132
; TYPE: DNA
; ORGANISM: human
US-60-248-505-387

alignment_scores:

Quality:	101.00	Length:	141
Ratio:	1.347	Gaps:	9
Percent Similarity:	53.191	Percent Identity:	32.624

alignment_block:

us-09-041-994-2_COPY_1018_1179 x us-60-248-505-387 ..

Align seg 1/1 to: us-60-248-505-387 from: 1 to: 5132

20 LeuValGlyProProSerAsnLeuGlnGlnGlnSerAspGluArgAla 36
||| |||
3353 CTCGGCGGCGCACCCAGGAACTGCAGCCGAGGTGT...GAGCGGAGGT 3399
36 uLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGlyLeuG 53
||| |||
3400 ACTCCCGGCGCTGGGTACGTCGCGGACCTGCACCCCACTCTCCACCG 3449
53 yGlnIleAspArgAlaLeuGlnGlyIleProGluLeuValAsn..... 66
|| |||
3450 AG.....CGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCT 3490
67 ..GlnGlnAlaLeuGln.....ProLysGlnAs 76
||| |||
3491 GCCAAGGCGCAGGCTCTACCTCTCCCTGCGTCTTGTCTCCCACTGAC 3540
76 pAlaPheGlnGlnGln.....GluAlaAlaValMetMetAspGlnLysAlaG 92
.....
3541 CCGCTCCCGCTCTCAGCCAGCAGGCTGCGCTGCGCTGCGCTGCGCTGCG 3590
92 y.....LeuTyrglyGlnThrTyrrProAlaGlnGlnGlyProPrometGln... 106
|| |||
3591 GGACATTTCTGGGCTCCAGCTTCTCCGAGCTCCCTCCACCGCAGCGA 3640
107 ..GlyIlePheHisLeuGlnGlnGlnSerProSerPheAsnSerMetMe 122
||| |||
3641 CCGGTGGCAGCTAT.....AGGAACCAAGGCTTCGGAAGCTTGC 3681
122 tAsnGlnMetAsnGlnGlnGlnGlnPheProLeuGlnGlnGlnGlnProAla 139
||| |||
3682 CAACCAGCAAAACAG.....ACCCGAGGAAGCCGCCCTC 3716
139 rgAlaAsnIleMetArgProArg 146
3717 CTCGCCCTTCCGCGAGCCCGG 3739

seq_name: /cgn2_6/prodata/2/pna/us60_NEW_COMB.seq:us-60-248-505-222

seq_documentation_block:

; Sequence 222, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: c1000918
; CURRENT APPLICATION NUMBER: US/60/248, 505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222
; LENGTH: 31818
; TYPE: DNA

ORGANISM: human
US-60-248-505-222

alignment_scores:
Quality: 101.00 Length: 141
Ratio: 1.347 Gaps: 9
Percent Similarity: 53.191 Percent Identity: 32.624

alignment_block:

US-09-041-994-2_COPY_1018_1179 x US-60-248-505-222 ..

Align seg 1/1 to: US-60-248-505-222 from: 1 to: 31818

```
20 LeuValGlyProProSerAsnLeuGlyGlnSerAspGluArgAla 36
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30039 CTCGGGGGGCCACGAGGAACTGACCCGAGGTGT...GACGGGAGGT 30085
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
36 uLeuAspGlnLeuHisThrLeuSerAsnThrAspAlaThrGlyLeu 53
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30086 ACTCCCGCGGCTGGTACGCTCCGGGACCTCCACCCCACTCCACCG 30135
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
53 LuGluIleAspArgAlaLeuGlyIleProGluLeuValAsn..... 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30136 AG.....CGGGCTCTCGGCTCGCGCTCTGGGCGAGCCAGGGGCT 30176
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67 ..GlnGlyGlnAlaLeuGlu.....ProLysGlnAs 76
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30177 GCCCAAGGCCAGGCTCTACCTCTCGGTTCCTTGTGTCCTCCACG 30226
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
76 palAPheGlnGlyGln..GluAlaAlaValMetLeuAspGlnLysAla 92
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30227 CCGCTCCCGCTCTACGACGACAGACTCGCTGGGCACTCGGAGAAAGCG 30276
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
92 Ly...LeuTyrGlyGlnThrTyrProAlaGlnGlyProPromeGln... 106
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30277 GGCACAAATTCCTTCCAGTCTTCCGGACTCCCTCTCCACGCGCA 30326
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
107 ..GlyGlyPheHisLeuGlnGlyGlnSerProSerPheAsnSerMet 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30327 CCGGGTGGAGCTAT.....AGGAAACAGGCTTCGGAAGTCTTGC 30367
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 LeAsnGlnMetAsnGlnGlnLysAsnPheProLeuGlnGlyMetHisP 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30368 CAACGACGACAAACAG.....ACCCGAGGAGCCCGCTC 30402
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
139 rGluAlaAsnIleMetArgProArg 146
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30403 CTCCTCGCTTCCGCGAGCCCGG 30425
```

seq_name: /cgn2_6/ptodata/2/pna/US09_NFM_COMB.seq:US-09-739-449-6007

seq_documentation_block:

```
; Sequence 6007, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 6007
; LENGTH: 2349
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-6007
```

alignment_scores:
Quality: 92.00 Length: 168

Ratio: 1.045 Gaps: 6
Percent Similarity: 52.381 Percent Identity: 24.405

alignment_block:

US-09-041-994-2_COPY_1018_1179 x US-09-739-449-6007 ..

Align seg 1/1 to: US-09-739-449-6007 from: 1 to: 2349

```
23 ProProSerAsnLeuGlyGlnSer.....As 32
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1375 CCCCAGCCAAACATGAACCAACCAAGCCGAGCTGCTGCCGCGCAGCGCA 1424
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
32 pGluArgAlaLeuAspGlnLeuHisThrLeu.....SerAsp 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1425 TCTGGAAACATGATGAACCAACATCGAAGACTTGGCCGTTCCGGCAATC 1474
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
47 hrAspAlaThrGly.....LeuGluGluIleAspArgAlaLeuGlyIle 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1475 GTACCGCGGCGGAGAAATCTCTCCAGCTACAGCC..... 1512
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
62 ProGluLeuValAsnGlnGlnAlaLeuGluProLysGlnAspAla 78
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1513 .....ATGATGAACAAATCTGACGCTGCGCGCCGCGCAGCGCAAGGCCA 1556
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
78 eGlnGlyGlnGluAlaAlaValMetLeuAspGlnLysAlaGlyLeuTyr 95
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1557 GCAGGACACCAACAGAGCGACAGATGCCGACAGATCGACACAGCTGG 1606
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
95 LysIleThrTyrProAlaGln.....GlyProPromeGlnGly 108
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1607 GCGAAATTCCTTACGACGACGACAGATGATGACGAGACTTCAAGCTC 1656
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
102 .....GlyProPromeGlnGly 108
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1657 GACCAAGCACTTCGAGACCGCATGCAAGCGCGGACCCACAGCAGGCGG 1706
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
108 yPheHisLeuGlnGlyGlnSerProSerPheAsnSerMetLeuGln 125
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1707 CGAAGCGGAAGACGACGACGATGGCGAGAAATGGCCACGCCACAGG 1756
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
125 eAsnGlnGlnGlnLysAsnPheProLeuGlnGlyMetHisProAlaAsn 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1757 AAGCGCAGCAAGTCAACCGGTCAACCAAGCCAGATGCTACAGAGCG 1806
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
142 IleMetArgProArgThrAsn...ThrProLysGlnLeuArgMetLeu 157
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1807 GGCACACAGCGCGAGCGACGATGACCGCGGAACAAATTCGCGCAGCGCT 1856
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
157 uGln 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1857 GAAA 1860
```

seq_name: /cgn2_6/ptodata/2/pna/US09_NFM_COMB.seq:US-09-739-449-217

seq_documentation_block:

```
; Sequence 217, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 217
; LENGTH: 397803
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(397803)
```

```
; OTHER INFORMATION: unsure at all n locations
US-09-739-449-217

alignment_scores:
    Quality: 92.00      Length: 168
    Ratio: 1.045      Gaps: 6
    Percent Similarity: 52.381      Percent Identity: 24.405

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-09-739-449-217 ..
Align seg 1/1 to: US-09-739-449-217 from: 1 to: 397803

23 ProProSerAsnLeuGlnGlyGlnSer.....As 32
   ||| .....|||.....|||.....|||
31580 CCCCAGCCCAACATGACACACCAACCCAGAACGTCGTGCCGACCGCGGA 31629

32 pGlnArgAlaLeuLeuAspGlnLeuHisThrLeu.....SerAsnT 47
   | : : : : : : : : : : : : : : : : : :
31630 TCTGGAATAACATGATGAAACAGATCGAAGACCTTCCCGTCCGCAATC 31679

47 hrAspAlaThrGly.....LeuGlnGluLeuAspArgAlaLeuGlyLe 61
   ||||| : : : : : : : : : : : : : : :
31680 GTGACGCGCGCCAGGAATGCTCTCGACGCTACAGCGC..... 31717

62 ProGluLeuValAsnGlnGlyGlnAlaLeuGluProLysGlnAspAlaPh 78
   : : : : : : : : : : : : : : : : : :
31718 .....ATGATGAACATCTGACAGCTGCGCGCCGCGCAGCGCAAGGCCA 31761

78 eGlnGlyGlnGlnAlaAlaValMetMetAspGlnLysAlaGlyLeuTyG 95
   ||||| : : : : : : : : : : : : : : :
31762 GCAAGGACGACAAACAGACGACAGCATGCGCGCAGATCGACAAAGCTGG 31811

95 LysLThrTyProAlaGln.....GlyProMetGlnGly 101
   ||| : : : |||
31812 GCGAATTCCTTCAGACGACGAGAAAGTTGATGACGACACTTTCAGCTC 31861

102 .....GlyProMetGlnGly 108
   ||| : : : |||
31862 GACCAACACTTCAGACGCGCATGCGACGCGCGCACGACGACGAGCGGG 31911

108 yPheHisLeuGlnGlnInsProSerPheAsnSerMetMetAsnGlnM 125
   : : : : : : : : : : : : : : : : : :
31912 CGAAGCGCAACAGCAGCAGAGTGGCGAGAAATGGCCAGCGCCACAGG 31961

125 etAsnGlnGlnGlyAsnProLeuGlnGlyMetHisProArgAlaAsn 141
   : : : : : : : : : : : : : : : : : :
31962 AAGGCGACAGATGACACCGGGTCAACAAGCCAGAAATGTCACGACGAG 32011

142 IleMetArgProArgTyThrAsn...ThrProLysGlnLeuArgMetGln 157
   : : : : : : : : : : : : : : : : : :
32012 GGACACACGCGGACGCGACCATGACCGGAAATTTGCGGACGCGCT 32061

157 uGln 158
   ||| : : : |||
32062 GAAA 32065

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-654-935A-99

seq_documentation_block:
; Sequence 99, Application US/09654935A
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Felyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qiong A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
```

```
; APPLICANT: Wang, Zhiwei
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/654,935A
; NUMBER OF SEQ ID NOS: 271
; SOFTWARE: pt_fl_genes Version 1.0
; SEQ ID NO 99
; LENGTH: 12279
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (8405)..(12226)
US-09-654-935A-99

alignment_scores:
    Quality: 86.00      Length: 126
    Ratio: 1.284      Gaps: 6
    Percent Similarity: 53.175      Percent Identity: 28.571

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-09-654-935A-99/rev ..
Align seg 1/1 to reverse of: US-09-654-935A-99 from: 1 to: 12279

2 GlnValSerHisGlyThrGlnAsnArgPro...LeuLeuArgAsnSerLe 17
   ||||| : : : : : : : : : : : : : : :
11884 CAGGTGCCACAGGTGGCCACCTGGCCCTGAACTCGACACGAGTGAGCT 11935

17 uAspAspLeuValGlyProProSerAsnLeuGlnGlyGlnSerAspGln 34
   : : : : : : : : : : : : : : : : : :
11834 GCAGATCTGCTGGCGCAT..... 11816

34 rGlnAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThr...AspAla 49
   : : : : : : : : : : : : : : : : : :
11815 ..GTAGCCCTGTGATGACCTTCACGAAATGCTGATCTGCTTGAACAG 11768

50 ThrGlyLeuGlnGluIleAspArgAlaLeuGlyIleProGluLeuValAs 66
   ||||| : : : : : : : : : : : : : : :
11767 CTGACGCTGACGGAACCTGCACAGAGCC.....GGCATGTGGCTCA 11727

66 nGlnGlyGlnAlaLeuGluProLysGlnAspAlaPhe..... 78
   : : : : : : : : : : : : : : : : : :
11726 GCAGGCTGTGGCTGTGAGTGAGTGAAAGCAGCTCTTGAAGGCCACATC 11677

79 .....GlnGlyGlnGlnAlaAlaValMetMetAsp 88
   : : : : : : : : : : : : : : : : : :
11676 ATGACCTTCAGCTGCAGCAGAGAGAGAAAGCGCGCTGTACTTGTCTAG 11627

89 GlnLysAlaGlyLeuTyGlnThrTyProAlaGln...GlyProP 104
   ||| : : : |||
11626 GCAAGCCCTGGGATGACAAATGTGAGAGCCCATGCTTGTACTTGA 11577

104 oMetGlnGlyGlyPheHisLeuGlnGly 113
   : : : : : : : : : : : : : : : : : :
11576 GCTCAGGCACTGCAGCATTCGCGGCG 11549

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12214

seq_documentation_block:
; Sequence 12214, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: Transcriptome
; CURRENT APPLICATION NUMBER: US/09/335,032
```

; CURRENT FILING DATE: 1999-06-16
 ; PRIOR APPLICATION NUMBER: US 60/035,917
 ; PRIOR FILING DATE: 1997-01-23
 ; PRIOR APPLICATION NUMBER: US 09/012,031
 ; PRIOR FILING DATE: 1998-01-22
 ; NUMBER OF SEQ ID NOS: 12219
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12214
 ; LENGTH: 666448
 ; TYPE: DNA
 ; ORGANISM: Saccharomyces cerevisiae
 ; US-09-335-032-12214

alignment_scores:
 Quality: 84.00 Length: 177
 Ratio: 0.884 Gaps: 8
 Percent Similarity: 53.672 Percent Identity: 25.989

alignment_block:

US-09-041-994-2_COPY_1018_1179 x US-09-335-032-12214 ..

Align seg 1/1 to: US-09-335-032-12214 from: 1 to: 666448

```

1 GluGlnValSerHisGlyThr..GlnAsnArgProLeuLeuArgAsnSe 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
365901 GAGCAGTTGTTTATACACCGCTCAGACAGTCTCCATACGAATAACA 365950
16 rLeuAspAspLeuValGlyProProSerAsnLeuGlnGlyL.. 30
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
365951 TACTAATGATTCGAAATGCAACGAAAGTGCACCGCTCAGCAGCTG 366000
31 ..SerAspGluArgAlaLeuLeuAspGlnLeuHisThr.....LeuLeu 44
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366001 CCTCTGAGTGGTGGTATTGCTTATAGAACTTCGTGTTACTA 366050
45 SerAsnThrAspAlaThrGlyLeuGlnGluLeuLeuAspAlaLeuGly 61
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366051 GCTCTCTACTGATGATCATATGGGAGACAGACAAAGATCAAGTTCAT 366100
61 eProGluLeuValAsnGln.....GlyGln. 69
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366101 ACCTTCTTTGCTAGCAATATCTCAATTTCTTTTACTAGTGCCAGC 366150
70 ..AlaLeuGlnProLysGlnAspAlaPheGlnGlyGlnAlaAlaVal 85
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366151 CTCAGCTTCCCCCAGCTCAGCAACATCGCAGCCACCAAGCACCCACAA 366200
86 MetMetAspGlnLysAlaGlyLeuTyGlyGlnThrTyProAlaGlnG 102
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366201 AAGTTGCACAGACAGACATCTACTCTGTAATTTTCCACACA 366250
102 yProProMetGlnGlyLysPheHisLeuGlnGln.....SerPro 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366251 GCAACCGTACACCAACCATGATATCATTTTGGCCAGTTCGCTGTAATG 366300
117 ePheAsnSerMetMetAsnGlnMetAsnGlnGlnGlyAsnPheProleu 133
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366301 GTTTTCATCTC.....AGACAAAGATTCAGTCCGCCAGC 366332
134 GlnGlnMetHisProArgAlaAsnIleMetArgProArgThrAsnThr.. 149
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366333 GAGGCTATGTCTCCAGTGCACCTGCCATGTTTACTAGCACATCTACAAA 366382
150 ..ProLysGlnLeuArgMetGlnLeuGlnGln 159
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366383 CCCCATAATGTTCACAGCAACACCAAGA 366413
seq_name: /cgn2_6/prcdat2/pna/us09_NEW_COMB.seq:us-09-801-833-6422
seq_documentation_block:
; Sequence 6422, Application US/09801833
; GENERAL INFORMATION:

```

; APPLICANT: Glucksmann, M. Alexandra
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
 ; FILE OF INVENTION: HUMAN BRAIN LIBRARY
 ; FILE REFERENCE: 1600.1037-005
 ; CURRENT APPLICATION NUMBER: US/09/801,833
 ; CURRENT FILING DATE: 2001-03-13
 ; PRIOR APPLICATION NUMBER: 09/371,168
 ; PRIOR FILING DATE: 1999-08-10
 ; PRIOR APPLICATION NUMBER: 60/095,907
 ; PRIOR FILING DATE: 1998-08-10
 ; PRIOR APPLICATION NUMBER: 60/103,145
 ; PRIOR FILING DATE: 1998-10-05
 ; NUMBER OF SEQ ID NOS: 8285
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6422
 ; LENGTH: 2055
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(2055)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-801-833-6422

alignment_scores:
 Quality: 81.00 Length: 157
 Ratio: 1.038 Gaps: 5
 Percent Similarity: 49.682 Percent Identity: 22.293

alignment_block:

US-09-041-994-2_COPY_1018_1179 x US-09-801-833-6422/rev ..

Align seg 1/1 to reverse of: US-09-801-833-6422 from: 1 to: 2055

```

22 GlyProProSerAsnLeuGluGlyGlnSerAspGluArgAlaLeuAs 38
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1025 GCGCGGCCAGGCCCCCAGGAAGGACCCAGCAGCAGCAGTACGTGA 976
38 pGlnLeuHisThrLeuLeuSerAsnThrAspAlaThrGly..... 51
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
975 TGAGTGCATGACAAAGATCCATGTTGGCGATGGCAGAGGGAGGCCAG 926
52 .....LeuGlnGluLeuAspArgAlaLeuGlyIleProGluLeuVal 65
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
925 GCCAGCCTGCGCTGCTGCTGCCCTGAGGGCCGCTGCGCAGTGCCTCG 876
66 AsnGlnGly.....GlnAlaLeuGluProLysGlnAspAlaPheGln 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
875 AACCGGCTGATGATCAGCAGCAGCTGTGTATGATGACCTCTACAGG 826
80 yGlnGluAlaAlaValMetMetAspGlnLysAlaGlyLeuTyGlyGln 97
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
825 TCAGCAGCAGCAGC.....CACAGTGAAG 803
97 hTyThrProAlaGlnGlyProProMetGlnGlyLysPheHisLeuGln 113
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
802 ATCACCCAGCCCCGCGGCGCTGTTCTTCGCGGAGAAAGTTGAGGCG 753
114 GlnSerProSerPheAsnSerMetMetAsnGlnMetAsnGln..... 127
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
752 AAGACGTCGAGCCCGCAGACAGAAAGACAGATGGCCAGAACACCCAA 703
128 .....GlnGlyAsnPheProLeuGlnGlyMetHis..ProArg 139
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
702 AGAGGAAGCCCGCAGAGCAGCAGTGAGAGTGGGCTTCCACACAG 653
140 AlaAsnIleMetArgProArgThrAsnThrProLysGlnLeuArgMetG 156
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
652 CCAACACAGCAGCAGAAAGAGGCCGAGGTCCCGCAGAAAGAAATACCTG 603
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
156 nLeuGlnGlnArgLeuGln 162
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```



```

103 ProhomMetGlnGlyPheHisLeuGlnGlnGlnSerProSerPheAs 119
1232 .....CAGGTACAGCCAGCAGGACATTTCACAGG 1260
119 nserMetMetAsnGlnMetAsnGlnGlnGlnGlnPheProLeuGlnGlym 136
1261 CCCAGGCGAGGTGAGCTGCAGCAGCAGGAGCAG...CCGCTGAAGCAG 1307
136 ethisProAlaGlnAlaAsnIleMetArgProArgThrAsnThr....Pro 150
1308 TGCAGCCACAG.....GTCCAGCCCGCAGCAGCATTCACAGCCCA 1348
151 LysGlnLeuArgMetGlnLeuGlnGlnArgLeuGln 162
1349 AGCAGGCTGACGCTGCAGCTGCAGACAGCAGGCTCAG 1384

seq_name: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:US-60-248-505-604

```

```

seq_documentation_block:
; Sequence 604, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: c1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 604
; LENGTH: 45333
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(45333)
; OTHER INFORMATION: n = A,T,C or G
US-60-248-505-604

```

```

alignment_scores:
Quality: 78.50 Length: 158
Ratio: 0.924 Gaps: 8
Percent Similarity: 53.797 Percent Identity: 25.949

```

```

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-60-248-505-604 ..

```

```

Align seg 1/1 to: US-60-248-505-604 from: 1 to: 45333

```

```

17 LeuAspAspLeuValGlyProProSerAsnLeuGlnGlnSerAspG1 33
1232 .....CAGGTACAGCCAGCAGGACATTTCACAGG 1260
20767 TTACAGACTCTTTCAGCAGCAGCAAAATTAGACAGCAGGTAATCA 20816
33 uArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAspAlaT 50
20817 CAGC.....CAGCAGGAGCATACAC.....AATGCCAATGCAG 20848
50 hrgLysLeuGlnGlnLeuAspArgAlaLeuGlyIleProGlnLeu.Valas 66
20849 TGCCTTTTACGCAAGTG.....AAAGTACCTCA 20877
66 nGlnGlyAlaLeuGlnLeuProLysGlnAspAlaPheGlnGlnGlnGly 82
20878 GAAACACACACTGCTACAGCAGCAGCAGCAGCCAGCAGCAGCAGCAGC 20927
83 .....AlaAlaValMetMet 87
20928 AGCAGCCGGTTTACACCTTCAGCCCGCAGCAGATATATGACACTTCAGCAG 20977
88 AspGlnLysAlaGlyLeuTyrglyGlnThrTyrglyProAlaGlnGlyPro 104

```

```

20978 CAGCAGCAGCAGCAGATCTCTACAGCAACCTTACCCCGCAGCAG...CCGCC 21024
104 ometGlnGlyPheHisLeuGlnGlnGlnGlnSerProSerPheAsnSer 121
21025 GCATCCATTTCACAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 21074
121 etMetAsnGlnMetAsnGlnGlnGlnGlnGlnPheProLeuGlnGlnGly 137
21075 AGTTTTCACAG.....CAACAGCTACAGTTTTCACAGCAACAGTTGCT 21118
138 ProArgAlaAsnIleMetArgProArgThrAsn...ThrProLysGlnLe 153
21119 CCTCCACAGCAGCAGCTGCATCGCCCTCAGCAGCAGAGCTCCAGCCCTTCA 21168
153 uArgMetGlnLeuGlnGlnArg 160
21169 GCAGCATGCTCCTGCAGCAGCAG 21190

seq_name: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:US-60-281-593-382

```

```

seq_documentation_block:
; Sequence 382, Application US/60281593
; GENERAL INFORMATION:
; APPLICANT: Amy Lasek
; TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER
; FILE REFERENCE: PA-0045 P
; CURRENT APPLICATION NUMBER: US/60/281,593
; CURRENT FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 519
; SOFTWARE: PERL Program
; SEQ ID NO 382
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 383376.7
; NAME/KEY: unsure
; LOCATION: 317
; OTHER INFORMATION: a, t, c, g, or other
US-60-281-593-382

```

```

alignment_scores:
Quality: 77.50 Length: 105
Ratio: 1.314 Gaps: 6
Percent Similarity: 56.190 Percent Identity: 30.476

```

```

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-60-281-593-382/rev ..

```

```

Align seg 1/1 to reverse of: US-60-281-593-382 from: 1 to: 551

```

```

62 ProGlnLeuValAsnGln...GlyGlnAlaLeuGlnLeuProLysGlnAspAl 77
1232 .....CAGGTACAGCCAGCAGGACATTTCACAGG 1260
364 CCAGAGTTAGTCCAAACAGCTGGGCGAGGCTGAGCCAGGCTCCCN. 317
77 aPheGlnGlyGlnGlnAlaAlaValMet.....MetAspG 89
316 .....TGGGAATGTGCTCTTCTCTCAGAGAGTGTTTCTAGAGC 277
89 InLysAlaGlyLeuTyrglyGlnThrTyrglyProAlaGlnGlyProPro... 104
276 TGAAGCCGCGCTTAACATGTAAATGCCACACAGAGGTAGGCGCCAGCAG 227
105 ...MetGlnGlyPheHisLeuGlnGlnGlnGlnGlnPheProLeuGlnGly 119
226 GGATGAGTGGGCTGCTCAGTACGAGCAGCCACCTCAACAGTTTGGGTCA 177
119 nserMetMetAsnGlnMetAsnGlnGlnGlnGlnGlnPheProLeuGlnGly 136
176 AATTCACCTGACTGAGCAGCCAGCAGCAAGCAGCTTGGAGAAAGCCAGGG 127

```

```
136 echIsProArq.....AlaasnlleMetArg 144
      |||||
126 CTCACCTCCCTCTCCTACGATGTTTGAGCTGAGCTCAGATACCTCCG 77
      |||||
145 ProArqThrasnThr 149
      |||||
76 CCAGAGAGCAGCAGCT 62

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-822-687-3
seq_documentation_block:
; Sequence 3, Application US/09822687
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria Alexandra
; TITLE OF INVENTION: 31939, A NOVEL HUMAN LEUCINE-RICH REPEAT
; FILE REFERENCE: 10448-038001
; CURRENT APPLICATION NUMBER: US/09/822,687
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,919
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2142
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-687-3

alignment_scores:
  Quality: 77.50      Length: 190
  Ratio: 0.881      Gaps: 9
  Percent Similarity: 46.316      Percent Identity: 24.737

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-09-822-687-3 ..
Align seg 1/1 to: US-09-822-687-3 from: 1 to: 2142

5 HtSLgYThrGlnAsnArqProLeuLeuArqAsnSerLeuAspLeuVa 21
  |||||
1614 CACGGCACC.....CGCCCCGGCTCTCGCGGCCACGAGAGAGCGCT 1657
  21 1.....GlyProPro 25
      |||||
1658 TCACGGTCCCATCACGATGTGACGAGAGACGCCCTCAAGGACCTGGAC 1707
  25 eRAsnleuGlnGlnSerAspGluArgAlaLeuLeuAspGlnLeuHis 41
  |||||
1708 GACGTATAGAGACCAAAATCATCATCGCTGCTT...CGTGCCCAT 1754
  42 ThrLeuSerAsnThrAspAlaThrGlyLeuGlnGluLeuAspArgAl 58
  |||||
1755 CACGTTCAATGCGCGGCTGCTGCTGCTTCTTCAAGCTGCGCAAC 1804
  58 aLeuGlyIleProGluLeuValAsnGlnGlyGlnAlaLeuGluProLys 75
  |||||
1805 ACCACACAGCTCCACAA.....GCACACAGGGGCCACGCGACCGT 1845
  75 lAspAlaIleheGlnGlnGlnGlnGlnAlaAlaValMetMetAspGlnLysAla 91
  |||||
1846 GAGATCATCATCAGCTGAGGAGAGAGCTGCCCGCGCTGCTGCTGCT 1895
  92 GluLeuThrGlyGln..... 96
  |||||
1896 GGGCGCGCGGGCGCGCTGCGCCAGTGGGGGTGCTGGCGGGGACACGC 1945
  97 .....ThyTrProAlaGlnGlnGlyProProMetG 106
  |||||
1946 ACCTGGCCCTGCCCGCTGAGAGACGACCTCAACACACACACACCTAC 1995
  |||||
```

```
106 lncGlyIlePheHisLeuGlnGlnSerProSerPheAsnSerMetMet 122
      |||||
1996 GT...GGTACGCGCTTCAAGGC.....GCATCAACAG 2024
      |||||
123 AsnGlnMetAsnGlnGlnGlnAsnProLeuGlnGlyMetHisProAr 139
      |||||
2025 CAGCAACCCACGAGGGGGGG.....CTCGGGGGCAAGGCCCGC 2065
      |||||
139 gaLaAsnIleMetArqProArqThrasnThrProLysGlnLeuArgMetG 156
      |||||
2066 CTGGCCTCAACTCCATCCATCAGAACCTCTGCT..... 2096
      |||||
156 lncGlnGlnArgLeuGln 162
2097 ..CTTCAAGAGCGCGCTCCAA 2114

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-822-687-1
seq_documentation_block:
; Sequence 1, Application US/09822687
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria Alexandra
; TITLE OF INVENTION: 31939, A NOVEL HUMAN LEUCINE-RICH REPEAT
; FILE REFERENCE: 10448-038001
; CURRENT APPLICATION NUMBER: US/09/822,687
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,919
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2493
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (187)...(2325)
US-09-822-687-1

alignment_scores:
  Quality: 77.50      Length: 190
  Ratio: 0.881      Gaps: 9
  Percent Similarity: 46.316      Percent Identity: 24.737

alignment_block:
US-09-041-994-2_COPY_1018_1179 x US-09-822-687-1 ..
Align seg 1/1 to: US-09-822-687-1 from: 1 to: 2493

5 HtSLgYThrGlnAsnArqProLeuLeuArqAsnSerLeuAspLeuVa 21
  |||||
1800 CACGGCACC.....CGCCCCGGCTCTCGCGGCCACGAGAGAGCGCT 1843
  21 1.....GlyProPro 25
      |||||
1844 TCACGGTCCCATCACGATGTGACGAGAGACGCCCTCAAGGACCTGGAC 1893
  25 eRAsnleuGlnGlnSerAspGluArgAlaLeuLeuAspGlnLeuHis 41
  |||||
1894 GACGTATAGAGACCAAAATCATCATCGCTGCTT...CGTGCCCAT 1940
  42 ThrLeuSerAsnThrAspAlaThrGlyLeuGlnGluLeuAspArgAl 58
  |||||
1941 CACGTTCAATGCGCGGCTGCTGCTGCTTCTTCAAGCTGCGCAAC 1990
  58 aLeuGlyIleProGluLeuValAsnGlnGlyGlnAlaLeuGluProLys 75
  |||||
1991 AGCACACAGCTCCACAA.....GCACACAGGGGCCACGCGACCGT 2031
  75 lAspAlaIleheGlnGlnGlnGlnGlnAlaAlaValMetMetAspGlnLysAla 91
  |||||
```

```
2032 GAGATCATCAACGTGAGAGACGAGCTGCCCGCGCTCGGCCGTGTCCGT 2081
92 GlyLeuTyrGlyGln..... 96
||| |||
2082 GGGCGCGCGCGCGCGCGTGGCCAGTGGGGGTGTGTGGCGGGGACACCC 2131
97 .....ThrTyrProIaGlnGlyProPomeTg 106
||| ||| |||
2132 ACCTGGGCCCTGCCGCCCTGTGAGGAGACACCTCAACACACACACTAC 2181
106 InGlyGlyPheHisLeuGlnGlyInSerProSerPheAsnSerMetMet 122
:: ||| ::||| |||
2182 GT...GGCTGACGGCTTCAAGGC.....GCACTACAG 2210
123 AsnGlnMetAsnGlnGlnGlyAsnPheProLeuGlnGlyMetHisProAr 139
::: ||| ::||| |||
2211 CAGCAACCCCGAGCGCGGGGG.....CTGCGGGGGCAAGGCCCGCC 2251
139 gAlaAsnIleMetArpProArGThrAsnThrProLysGlnLeuArGmetG 156
::: ||| ::||| |||
2252 CTGGCCCTCAACTCCATCCACGACACCTCTGCT..... 2282
156 InLeuGlnGlnArGLeuGln 162
||| ||| ||| ||| |||
2283 ..CTTCAAGAGCGGCTCCAA 2300
```

This Page Blank (uspto)

Align seg 1/1 to: US-07-667-276A-3 from: 1 to: 2017

```

4 LysGlyHisLysLysLeuLeuGlnLeuLeuThrCysSerSerAspAsp 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
505 AAGGTAAACAAGAGAGAGTAAAGCTTCCAAACAAGCCAAAGAAAA 554
20 gclHisSerSerLeuThrAsnSerProLeuAspSerCysLysGln 37
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
555 AGCT...AAGGCCGTCTCTCCCTCTCCGAAATCTCATCTCATCT 601
37 eSerSerValSerValThrSerProSerGlyValSerSerThrSer 53
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
602 CATCTTCATCTGAATCTGAATCTGAGTCTGAGTCTGAATCTGAAT 651
54 GlyValSerSerThrSerAsnMetHisGlySerLeuLeuGlnLysH 70
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
652 TCTTCATCTCATCTCCCTCATCTCATCTCTCTCTCTCTCTCTCT 684
70 sArgIleuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluVal 87
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
685 .....TCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 712
87 lAlaSerIleThrAlaGlnAlaThrGlyAspThrSerSerIleThr 103
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
713 AAGCTGAACCAAGAGAGAGAGATCCAGAGATCTCTCTCTCTCTCT 762
104 CysGlyAspGlyAsnValValLysGlnGlnGlnLeuSerProLys 120
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
763 GACTCTTCTCCGACGACAGAGAGAGAGAGAGAGAGAGAGAGAG 812
120 sGluAsnAsn 123
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
813 GGAGAGATCA 822

```

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-209-747-1

```

seq_documentation_block:
: Sequence 1, Application US/08209747
: Patent No. 5733771
: GENERAL INFORMATION:
: APPLICANT: Lewis, Randolph V.
: TITLE OF INVENTION: CDNA Encoding Minor Ampullate Spider
: TITLE OF INVENTION: Silk Proteins
: NUMBER OF SEQUENCES: 56
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Birch, Stewart, Kolasch & Birch
: STREET: P.O. Box 747
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: USA
: ZIP: 22040-3487
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/209/747
: FILING DATE: 14-MAR-1994
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Murphy Jr., Gerald M.
: REGISTRATION NUMBER: 28,977
: REFERENCE/DOCKET NUMBER: 1447-104P
: TELEPHONE: 703-205-8000
: TELEFAX: 703-205-8050
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2793 base pairs
: TYPE: nucleic acid

```

```

: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: Nephila clavipes
: TISSUE TYPE: minor ampullate gland
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 183..2675
: OTHER INFORMATION: /product="N. clavipes minor
: OTHER INFORMATION: ampullate silk protein"
US-08-209-747-1

```

```

alignment_scores:
Quality: 83.50 length: 116
Ratio: 1.176 Gaps: 4
Percent Similarity: 61.207 Percent Identity: 31.897

```

alignment_block:

US-09-041-994-2_copy_613_752 x US-08-209-747-1/rev ..

Align seg 1/1 to reverse of: US-08-209-747-1 from: 1 to: 2793

```

14 ThrCysSerSerAspAspArgGlyHisSerSer..... 24
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
860 TCTGTCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 811
25 .....LeuThrAsnSerProLeuAspSerSerCysLysGlnSer 39
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
810 ACCCACCCTTGACCCGCTAGCCTCCAGCTCCAGCAGCAGCAGCAG 761
39 aSerValThrSerProSerGlyValSerSerSerThrSerGlyVal 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
760 GCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 711
56 SerSerThrSerAsnMetHisGlySerLeuGln..... 67
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
710 ACCAGCAGCAGCAGCAGCCG.CACGAGCTCTGCTCCGACCATATCC 662
68 .....GluLysHisArgIleLeuHisLysLeuGln.Asn 79
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
661 TGTCACCATATTCACGACACCGCGCGCTGCACCTGCCACGACCC 612
80 GlyAsnSerProAlaGluValAlaLysIleThrAlaGlnAlaThr 96
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
611 ACCAGCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 562
96 sAspThrSerSerIle.....ThrSerCysGlyAspGlyAsnVal 109
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
561 CTGCTCCAGCAGCTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAT 516
seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-458-298-1
seq_documentation_block:
: Sequence 1, Application US/08458298
: Patent No. 5756677
: GENERAL INFORMATION:
: APPLICANT: Lewis, Randolph V.
: TITLE OF INVENTION: CDNA Encoding Minor Ampullate Spider
: TITLE OF INVENTION: Silk Proteins
: NUMBER OF SEQUENCES: 56
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Birch, Stewart, Kolasch & Birch
: STREET: P.O. Box 747
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: USA
: ZIP: 22040-3487
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk

```

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/458, 298
  FILING DATE: 02-JUN-1995
  CLASSIFICATION: 530
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/209,747
    FILING DATE: 14-MAR-1994
  ATTORNEY/AGENT INFORMATION:
    NAME: Murphy Jr., Gerald M.
    REGISTRATION NUMBER: 28,977
  REFERENCE/DOCKET NUMBER: 1447-104P
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 703-205-8000
    TELEFAX: 703-205-8050
  INFORMATION FOR SEQ ID NO: 1:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 2793 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
      MOLECULE TYPE: cDNA
      HYPOTHETICAL: NO
      ORIGINAL SOURCE:
        ORGANISM: Nephila clavipes
        TISSUE TYPE: minor ampullate gland
      FEATURE:
        NAME/KEY: CDS
        LOCATION: 183..2675
        OTHER INFORMATION: /product= "N. clavipes minor
        OTHER INFORMATION: ampullate silk protein"
US-08-458-298-1
```

```
alignment_scores:
  Quality: 83.50      Length: 116
  Ratio: 1.176      Gaps: 4
  Percent Similarity: 61.207      Percent Identity: 31.897
```

alignment_block:

US-09-041-994-2_COPY_613_752 x US-08-458-298-1/rev ..

Align seg 1/1 to reverse of: US-08-458-298-1 from: 1 to: 2793

```
14 ThrCysSerSerAspArgGlyHisSerSer..... 24
   ::::::::::::::::::::
860 TCCTGCTCCAGCAGCAGCAGCAGCAGCAGCTGCTCCGCGCAGCCT 811
25 .....LeuThrAsnSerProLeuAspSerSerCysLysGluSerSerV 39
   ::::::::::::::::::::
810 ACCCACCCTTGACACCGTAGCTCCAGCTCCAGCCAGCCGCGGCTCCT 761
39 alSerValThrSerProSerGlyValSerSerSerThrSerGlyGlyVal 55
   ||| ||||| ||| ||| ::::::::::::::::::::
760 GCTCCAGCAGCAGCAGCAGCAGCAGCTTACCGTAAACCGCAGCAGCTCTGCTCCT 711
56 SerSerThrSerAsnMetHisGlySerLeuLeuGln..... 67
   ::::::::::::::::::::
710 ACCAGCAGCAGCAGCAGCAGCAGCAGCTCTGCTCCGCGCAGCATATCCGCGCT 662
68 .....GluLysHisArgGlyLeuHisLysLeuLeuGln.Asn 79
   ::::::::::::::::::::
661 TCTCCACCATATTCACGACACCGCGGCTGCACACCGCCCGCCAGCAGCAGC 612
80 GlysAsnSerProAlaGluValAlaLysIleThrAlaGlnAlaThrGlyVal 96
   ::::::::::::::::::::
611 ACCAGCTCCAGCAGCAGCAGCAGCAGCTTACCGTAAACCGCAGCAGCTCTGCTCCT 562
96 saspthrSerSerIle.....ThrSerCysGlyAspGlyAsnVal 109
   ::::::::::::::::::::
561 CTGCTCCAGCAGCAGCAGCAGCAGCAGCAGCTCTGCTCCGCGCAGCAGCATA 516
```

seq_name: /cgn2_6/ptodata/1/ina/5A.COMB.seq:US-08-320-559-29

seq_documentation_block:

```
: Sequence 29, Application US/08320559
: Patent No. 5633135
: GENERAL INFORMATION:
: APPLICANT: Croce, Carlo
: APPLICANT: Canaan, Eli
: TITLE OF INVENTION: Diagnostics, Therapeutics and Methods for
: TITLE OF INVENTION: Detection and Treatment of Acute Leukemias
: TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the
: NUMBER OF SEQUENCES: 44
: CORRESPONDENCE ADDRESS:
: ADDRESS: Woodcock Washburn Kurtz Mackiewicz & NO. 5633135rls
: STREET: One Liberty Place - 46th Floor
: CITY: Philadelphia
: STATE: PA
: COUNTRY: USA
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
: COMPUTER: IBM PS/2
: OPERATING SYSTEM: PC-DOS
: SOFTWARE: WORDPERFECT 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/320,559
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/062,443
: FILING DATE: 14 MAY 1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/971,094
: FILING DATE: 30-OCT-92
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/888,830
: FILING DATE: 27-MAY-92
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/805,093
: FILING DATE: 11-DEC-91
: ATTORNEY/AGENT INFORMATION:
: NAME: DeLuca, Mark
: REGISTRATION NUMBER: 33,229
: REFERENCE/DOCKET NUMBER: TJU-0855
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (215) 568-3100
: TELEFAX: (215) 568-3439
: INFORMATION FOR SEQ ID NO: 29:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 3376 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 196..1902
US-08-320-559-29
```

```
alignment_scores:
  Quality: 82.00      Length: 170
  Ratio: 1.000      Gaps: 5
  Percent Similarity: 48.235      Percent Identity: 24.706
```

alignment_block:

US-09-041-994-2_COPY_613_752 x US-08-320-559-29 ..

Align seg 1/1 to: US-08-320-559-29 from: 1 to: 3376

7 LysLysLeuLeuGlnLeu.....LeuThrCysSe 16

```
592 AGAAGTTGCTGAAGGCGAGGGGACCCCTTAATAGAGTATTCATTACCAG 641
16 rSerAspAspArgLysHisSerSerLeuThrAsnSerProLeuAspSerS 33
442 CAGACGACGACGACGACGACGATAGCAGACGACGACGACGACGACGAC 691
33 erCysLysGlnSerSerValSerValThrSerProSerGlyValSerSer 49
692 GCAGTACGACGACGACGACGACGACGACGACGACGACGACGACGACG 741
50 SerThrSerGlyGlyValSerSerThr..... 58
742 AGCAGCAGCAGCAGCAGCAGTAGTACCAAGTTTTCACAAAGCTCACAATT 791
59 .....SerAsnMetHisGlys 64
792 AATGAGAGGACACAGAGAAAACCTTCTAAGAGACTCCAGAGAACATATAA 841
64 erLeuLeuGlnGlnLysHisArgLysHisLysLysLeuLeuGlnAsnGly 80
842 GTGCCTTCAAGAAACCTTCCAGAGGATCACAACAAATCTTCCAAAGAAATCC 891
81 .....AsnSerProAlaGlnValAlaLysLysIleThrAl 91
892 TCTAAGAAACCCAAAGAAATATAACCTGAAAGAGAGAAATATGTTCC 941
91 agLnaIleThrGlyLysAspThrSer..... 99
942 TAAGATGGCCCTCAAGAGACCTTAACCACTGCAAAAGAGCCAAACACAG 991
100 .....SerIleThrSerCysGlyAspGlyAsn..... 108
992 ATACTAATTACTTCACCATCAGTGACAGACAAAGATAGAGGCTCTAGT 1041
109 .....ValValLysGlnGlnGlnLeuSerProLysLysGly 121
1042 AAAAGCCCCCATTTTCAGATTCTGAAGAACTCTACGCCAATAAAGGAA 1091
121 uAsnAsnAla 124
1092 AAAGAGTAGC 1101

seq_name: /cgn2.6/ptodata/1/lna/6A_COMB.seq:US-08-545-860D-29
seq_documentation_block:
; Sequence 29, Application US/08545860D
; Patent No. 6040140
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; NUMBER OF INVENTION: Resulting from Chromosome Abnormalities in the All-1 Region
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESSEE: No. 6040140ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,860D
; FILING DATE: 07-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```

```
APPLICATION NUMBER: PCT/US94/04496
FILING DATE: 22-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/10930
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/327,392
FILING DATE: 19-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/320,559
FILING DATE: 11-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,443
FILING DATE: 14-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/971,094
FILING DATE: 30-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/888,839
FILING DATE: 27-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/805,093
FILING DATE: 11-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Deluca Esq., Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-1262
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3439
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 3376 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 196..1902
US-08-545-860D-29

alignment_scores:
Quality: 82.00 Length: 170
Ratio: 1.000 Gaps: 5
Percent Similarity: 48.235 Percent Identity: 24.706

alignment_block:
US-09-041-994-2_COPY_613_752 x US-08-545-860D-29 ..
Align seg 1/1 to: US-08-545-860D-29 from: 1 to: 3376
7 LysLysLeuGlnGlnLeu.....LeuThrCysSe 16
592 AGAAGTTGCTGAAGGCGAGGGGACCCCTTAATAGAGTATTCATTACCAG 641
16 rSerAspAspArgLysHisSerSerLeuThrAsnSerProLeuAspSerS 33
442 CAGACGACGACGACGACGACGATAGCAGACGACGACGACGACGACGAC 691
33 erCysLysGlnSerSerValSerValThrSerProSerGlyValSerSer 49
692 GCAGTACGACGACGACGACGACGACGACGACGACGACGACGACGACG 741
50 SerThrSerGlyGlyValSerSerThr.....SerAsnMetHisGlys 64
742 AGCAGCAGCAGCAGCAGCAGTAGTACCAAGTTTTCACAAAGCTCACAATT 791
59 .....AATGAGAGGACACAGAGAAAACCTTCTAAGAGACTCCAGAGAACATATAA 841
```

```

64   erleuleugIngluLysHisArgIleLeuHisLysLeuGlnasnGly 80
    || .....||| ||| .....||| .....|||
842  GTGGCTTCAAGAACCTTCAGGATGACACAAATCTCCAAAGATCC 891
81   .....AsnSerProAlaGluValAlaLysIleThrAl 91
    |||.....|||.....|||.....|||
892  TCTAAGAACCCAAAGAAATTAACCACTGAAAGAGAAATAGTTCC 941
91   aglnAlaThrGlyLysAspThrSer..... 99
    .....|||.....|||.....|||
942  TAAGATGGCTTCAAGAACCTTAACCCATGTCAAAAGCCAAACAC 991
100   .....SerIleThrSerCysGlyAspGlyAsn..... 108
    .....|||.....|||.....|||
992  ATAGTAAGTACTACCATCACCATCAGCTGACAGATTAAGAAAGCTCTAGT 1041
109   .....ValValLysGlnGluGlnLeuSerProLysLysGly 121
    .....|||.....|||.....|||
1042  AAAAGCCCCCCTTCAGATTTCTGAAGAACTCTCAGCCAAAGAAAGAA 1091
121  uasnAsnAla 124
    .....|||.....|||.....|||
1092  AAAGAGTAGC 1101
    .....|||.....|||.....|||

seq_name: /cgn2_6/ptodata/1/lna/PCTUS_COMB.seq:PCT-US94-04496-29

seq_documentation_block:
: Sequence 29, Application PC/TUS9404496
: GENERAL INFORMATION:
: APPLICANT: Croce, Carlo
: APPLICANT: Canaan, Eli
: TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
: TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
: NUMBER OF SEQUENCES: 86
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Woodcock, Washburn, Kurtz, Macklewitz &
: STREET: One Liberty Place, 46th floor
: CITY: Philadelphia
: STATE: Pennsylvania
: COUNTRY: USA
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US94/04496
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Deluca Esq., Mark
: REGISTRATION NUMBER: 33, 229
: REFERENCE/DOCKET NUMBER: TJU-1242
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (215) 568-3439
: TELEFAX: (215) 568-3100
: INFORMATION FOR SEQ ID NO: 29:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3376 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 196..1902
: PCT-US94-04496-29

```

```

Quality: 82.00 Length: 170
Ratio: 1.000 Gaps: 5
Percent Similarity: 48.235 Percent Identity: 24.706

alignment_block:
US-09-041-994-2_COPY_613_752 x PCT-US94-04496-29 ..
Align seg 1/1 to: PCT-US94-04496-29 from: 1 to: 3376

7 LysLysLeuLeuGlnLeu.....LeuThrCysSe 16
    .....|||.....|||.....|||
592  AGAAGTTGCTGAAGGACGAGAGGACCTTAATAGATATTCATACAG 641
16   rSerAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSer 33
    .....|||.....|||.....|||
642  CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 691
33   erCysLysGluSerSerValSerValThrSerProSerGlyValSerSer 49
    .....|||.....|||.....|||
692  GCAGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 741
50   SerThrSerGlyValSerSerThr.....SerAsnMetHisGly 58
    .....|||.....|||.....|||
742  AGCAGCAGCAGCAGCAGCAGCAGTATGATCCAGTTTCAAAAGCCTCACAAT 791
59   .....SerAsnMetHisGly 64
792  AATGAGAGACGACAAAGAAACCTTGAAGACTCCAGAAATATATAA 841
64   erleuleugIngluLysHisArgIleLeuHisLysLeuGlnasnGly 80
    || .....||| ||| .....||| .....|||
842  GTGGCTTCAAGAACCTTCAGGATGACACAAATCTCCAAAGATCC 891
81   .....AsnSerProAlaGluValAlaLysIleThrAl 91
    |||.....|||.....|||.....|||
892  TCTAAGAACCCAAAGAAATTAACCACTGAAAGAGAAATAGTTCC 941
91   aglnAlaThrGlyLysAspThrSer..... 99
    .....|||.....|||.....|||
942  TAAGATGGCTTCAAGAACCTTAACCCATGTCAAAAGCCAAACAC 991
100   .....SerIleThrSerCysGlyAspGlyAsn..... 108
    .....|||.....|||.....|||
992  ATAGTAAGTACTACCATCACCATCAGCTGACAGATTAAGAAAGCTCTAGT 1041
109   .....ValValLysGlnGluGlnLeuSerProLysLysGly 121
    .....|||.....|||.....|||
1042  AAAAGCCCCCCTTCAGATTTCTGAAGAACTCTCAGCCAAAGAAAGAA 1091
121  uasnAsnAla 124
    .....|||.....|||.....|||
1092  AAAGAGTAGC 1101
    .....|||.....|||.....|||

seq_name: /cgn2_6/ptodata/1/lna/5B_COMB.seq:US-08-728-323A-1

seq_documentation_block:
: Sequence 1, Application US/08728323A
: Patent No. 5948676
: GENERAL INFORMATION:
: APPLICANT: Chang, Yuan
: APPLICANT: Bohenzky, Roy A.
: APPLICANT: Russo, James J.
: APPLICANT: Edelman, Isidore S.
: APPLICANT: Moore, Patrick S.
: TITLE OF INVENTION: Immediate Early Protein From Kaposi's
: TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Cooper & Dunham LLP
: STREET: 1185 Avenue of the Americas
: CITY: New York
: STATE: New York

```

alignment_scores:


```
91 laginalathcglylyasphrthserlethrsercysgly 105
      ::::::::::::::::::::|||
316 GCACGACGACGACGACGACGACGACGACGACGACGCG 359
seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:US-08-864-038A-1

seq_documentation_block:
; Sequence 1, Application US/08864038A
; Patent No. 6001592
; GENERAL INFORMATION:
; APPLICANT: Kunio NAKASHIMA et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
; TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
; TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
; TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 812-5 Hirano
; STREET: Ishinden
; CITY: Tsu-city
; STATE: Mie-prefecture
; COUNTRY: JAPAN
; ZIP: 514-01
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; OPERATING SYSTEM: Microsoft Windows 95
; SOFTWARE: Word Perfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,038A
; FILING DATE: May 28, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-184459
; FILING DATE: 15-July-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-5610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)986-2340
; TELEFAX: (212)953-7733
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2214
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: Pinctada fucata
; CELL TYPE: mantle epithelial cell
; US-08-864-038A-1

alignment_scores:
Quality: 78.50 Length: 84
Ratio: 1.539 Gaps: 2
Percent Similarity: 60.714 Percent Identity: 29.762

alignment_block:
US-09-041-994-2_COPY_613_752 x US-08-864-038A-1/rev ..
Align seg 1/1 to reverse of: US-08-864-038A-1 from: 1 to: 2214

16 SerSerAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSe 32
   ::::::::::::::::::::|||
471 ACCTCGAGACCTCGCCCAATCCAGTCCGAGTCCAGCACCTGCTCTG 422
32 rSerCysIysGluSerSerValThrSerProSerGlyValSers 49
   ||||| ||||| ::::::::::::::::::::|||
421 CTCGCGCGCGCTCGTCACGACGACGACGACGACGACGACGACCA 372
49 erSerThrSerGlyGlyValSerSerThrSerAsnMet.HtsGlySerL 65
```

```
371 CCACGACGACCTCGGCGCGAGTCCATTCACCGCATCATCATCCAGTC 322
      ::::::::::::::::::::|||
65 euleuGlnGlyHisArgIleLeuHis..... 74
321 CCATTGATCATGTCATGCTGCTGCTATGTCATATTCACCATCTC 272
75 .....LysLeuLeuGlnAsnGlyAsnSerPr 83
271 CACGCGCGCGCGCGCTCGGCTGCTCTCCAGCACGCGCGCTGCC 222
83 o 83
221 C 221

seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:US-08-864-038A-2

seq_documentation_block:
; Sequence 2, Application US/08864038A
; Patent No. 6001592
; GENERAL INFORMATION:
; APPLICANT: Kunio NAKASHIMA et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
; TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
; TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
; TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBOD
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 812-5 Hirano
; STREET: Ishinden
; CITY: Tsu-city
; STATE: Mie-prefecture
; COUNTRY: JAPAN
; ZIP: 514-01
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; OPERATING SYSTEM: Microsoft Windows 95
; SOFTWARE: Word Perfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,038A
; FILING DATE: May 28, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-184459
; FILING DATE: 15-July-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-5610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)986-2340
; TELEFAX: (212)953-7733
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3331
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Pinctada fucata
; CELL TYPE: mantle epithelial cell
; FEATURE: mRNA
; LOCATION: from 1 to 3331
; IDENTIFICATION METHOD: E (by experiment)
; US-08-864-038A-2

alignment_scores:
Quality: 78.50 Length: 84
Ratio: 1.539 Gaps: 2
Percent Similarity: 60.714 Percent Identity: 29.762
```


alignment_block:

US-09-041-994-2_COPY_613_752 x US-08-864-038A-2/rev ..

Align seg 1/1 to reverse of: US-08-864-038A-2 from: 1 to: 3331

```

16 SerSerAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSe 32
   ::::::::::: ::::::::::: :::::::::::
520 ACCGCCGAGACCTCCGCCCATCCAGTCGACGACGACCTGCTCTG 471
   ::::::::::: ::::::::::: :::::::::::
32 rSerCysLysGluSerSerValSerValThrSerProSerGlyValSer 49
   ::::::::::: ::::::::::: :::::::::::
470 CTCTCGCGCCCTGCTCCAGACACGACGACGACCTCCGCGACACCA 421
   ::::::::::: ::::::::::: :::::::::::
49 erserThrSerGlyGlyValSerSerThrSerAsnMet..HisGlySerL 65
   ::::::::::: ::::::::::: :::::::::::
420 CCAGCACCACCTCCGCGCGCGCTCCCATTCACCGCTCATCCCACTC 371
   ::::::::::: ::::::::::: :::::::::::
65 euleuGlnGluLysHisArgIleLeuHis..... 74
   ::::::::::: :::::::::::
370 CCATTCATCATCGTCATCGCTGCTCATCTAGTCATATTCACCAATCTC 321
   ::::::::::: :::::::::::
75 .....LysLeuLeuGlnAsnGlyAsnSerPr 83
   ::::::::::: :::::::::::
320 CACCGCGCGCGCGCGCTCCGCGCTCCTCCAGCAGCCGCGCTCC 271
   ::::::::::: :::::::::::
83 o 83
   |
270 C 270

```

seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:US-08-864-038A-4

seq_documentation_block:

Sequence 4, Application US/08864038A

Patent No. 6001592

GENERAL INFORMATION:

APPLICANT: Kunio NAKASHIMA et al.

TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR

TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID

TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING

TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: 812-5 Hirano

STREET: Isshinden

CITY: Tsu-city

STATE: Mie-prefecture

COUNTRY: JAPAN

ZIP: 514-01

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: Microsoft Windows 95

SOFTWARE: Word Perfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/864, 038A

FILING DATE: May 28, 1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 8-184459

FILING DATE: 15-July-1996

ATTORNEY/AGENT INFORMATION:

NAME: C. Bruce Hamburg

REGISTRATION NUMBER: 22,389

REFERENCE/DOCKET NUMBER: F-5610

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212)986-2340

TELEFAX: (212)953-7733

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 3331

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA to mRNA

ORIGINAL SOURCE: Pinctada fucata

ORGANISM: Pinctada fucata

CELL TYPE: mantle epithelial cell

FEATURE:

NAME/KEY: CDS

LOCATION: from 50 to 2263

IDENTIFICATION METHOD: P (by similarity to some other pattern)

US-08-864-038A-4

alignment_scores:

Quality:	Ratio:	Length:
78.50	1.539	84
Percent Similarity: 60.714	Percent Identity: 29.762	

alignment_block:

US-09-041-994-2_COPY_613_752 x US-08-864-038A-4/rev ..

Align seg 1/1 to reverse of: US-08-864-038A-4 from: 1 to: 3331

```

16 SerSerAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSe 32
   ::::::::::: ::::::::::: :::::::::::
520 ACCGCCGAGACCTCCGCCCATCCAGTCGACGACGACCTGCTCTG 471
   ::::::::::: ::::::::::: :::::::::::
32 rSerCysLysGluSerSerValSerValThrSerProSerGlyValSer 49
   ::::::::::: ::::::::::: :::::::::::
470 CTCTCGCGCCCTGCTCCAGACACGACGACGACCTCCGCGACACCA 421
   ::::::::::: ::::::::::: :::::::::::
49 erserThrSerGlyGlyValSerSerThrSerAsnMet..HisGlySerL 65
   ::::::::::: ::::::::::: :::::::::::
420 CCAGCACCACCTCCGCGCGCGCTCCCATTCACCGCTCATCCCACTC 371
   ::::::::::: ::::::::::: :::::::::::
65 euleuGlnGluLysHisArgIleLeuHis..... 74
   ::::::::::: :::::::::::
370 CCATTCATCATCGTCATCGCTGCTCATCTAGTCATATTCACCAATCTC 321
   ::::::::::: :::::::::::
75 .....LysLeuLeuGlnAsnGlyAsnSerPr 83
   ::::::::::: :::::::::::
320 CACCGCGCGCGCGCGCTCCGCGCTCCTCCAGCAGCCGCGCTCC 271
   ::::::::::: :::::::::::
83 o 83
   |
270 C 270

```

seq_name: /cgn2_6/ptodata/1/lna/5A_COMB.seq:US-08-072-610-1

seq_documentation_block:

Sequence 1, Application US/08072610

Patent No. 5532133

GENERAL INFORMATION:

APPLICANT: Barnwell, John

TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens,

TITLE OF INVENTION: Monoclonal Antibodies, and Diagnostic Assays

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Darby and Darby

STREET: 805 Third Ave.

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10022-7513

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/072, 610

FILING DATE: 19930602

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

```
NAME: Gogoris, Adda
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 5986/07686
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)527-7700
TELEFAX: (212)753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3337 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
MOLECULE TYPE: linear
TOPOLOGY: linear
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium vivax
IMMEDIATE SOURCE:
CLONE: pVMB3.3.1
US-08-072-610-1
```

```
alignment_scores:
Quality: 78.50 Length: 94
Ratio: 1.377 Gaps: 2
Percent Similarity: 60.638 Percent Identity: 30.851
```

alignment_block:

US-09-041-994-2_COPY_613_752 x US-08-072-610-1/rev ..

Align seg 1/1 to reverse of: US-08-072-610-1 from: 1 to: 3337

```
14 ThrCysSerSerAspArgGlyHisSerSerLeuThrAsnSerProle 30
||| ||||| ||| ||||| |||||
2368 ACCTCTTCTTCACTACCGCTGCTACTTCTACTACCGCTGCTAC 2319
30 uAspSerSerCysLysGlnSerSerValThrSerProSerGlyV 47
||||| ||||| ||||| ||||| |||||
2318 CTTCTACTTCT.....TCTACTTCTTCCACTTCTCTCTCTCTG 2278
47 aLSerSerSerThrSerGlyValSerSerThrSerAsnMetHisGly 63
||||| ||||| ||||| ||||| |||||
2277 GACCTCTTCCACTTCTCTGCTGCTACTTCTTCCACTTCTCTGCT 2228
64 SerLeuLeuGlnGlyHisArgIleLeuHisLysLeuLeuGlnAsnI 80
||| ||||| ||||| ||||| |||||
2227 TCC.....ACTTCTCTCTG 2214
80 yAsnSerProAlaGluValAlaLysIleThrAlaGlnAlaThrGlyLysA 97
||||| ||||| ||||| ||||| |||||
2213 TACCTCTTCCACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2164
97 sPThrSerSerIleThrSerCysGlyAspGly 107
||||| ||||| ||||| ||||| |||||
2163 GACCTCTTCCACTTCTTCTACTTCTGCTGCTGCT 2132
```

seq_name: /cgn2_6/ptodata/1/lna/5B_COMB.seq:US-08-719-822B-1

seq_documentation_block:

```
; Sequence 1, Application US/08719822B
; Patent No. 5874527
; GENERAL INFORMATION:
; APPLICANT: Barnwell, John
; TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby and Darby
; STREET: 805 Third Ave.
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022-7513
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/719,822B
FILING DATE: 09/30/96
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gogoris, Adda
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 5986/17686US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)527-7700
TELEFAX: (212)753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3337 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
TOPOLOGY: linear
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium vivax
IMMEDIATE SOURCE:
CLONE: pVMB3.3.1
US-08-719-822B-1
```

```
alignment_scores:
Quality: 78.50 Length: 94
Ratio: 1.377 Gaps: 2
Percent Similarity: 60.638 Percent Identity: 30.851
```

alignment_block:

US-09-041-994-2_COPY_613_752 x US-08-719-822B-1/rev ..

Align seg 1/1 to reverse of: US-08-719-822B-1 from: 1 to: 3337

```
14 ThrCysSerSerAspArgGlyHisSerSerLeuThrAsnSerProle 30
||| ||||| ||| ||||| |||||
2368 ACCTCTTCTTCACTACCGCTGCTACTTCTACTTCTACTACCGCTGCTAC 2319
30 uAspSerSerCysLysGlnSerSerValThrSerProSerGlyV 47
||||| ||||| ||||| ||||| |||||
2318 CTTCTACTTCT.....TCTACTTCTTCCACTTCTCTCTCTCTCTG 2278
47 aLSerSerSerThrSerGlyValSerSerThrSerAsnMetHisGly 63
||||| ||||| ||||| ||||| |||||
2277 GACCTCTTCCACTTCTCTGCTGCTACTTCTTCCACTTCTCTGCTACTCT 2228
64 SerLeuLeuGlnGlyHisArgIleLeuHisLysLeuLeuGlnAsnI 80
||| ||||| ||||| ||||| |||||
2227 TCC.....ACTTCTCTCTG 2214
80 yAsnSerProAlaGluValAlaLysIleThrAlaGlnAlaThrGlyLysA 97
||||| ||||| ||||| ||||| |||||
2213 TACCTCTTCCACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2164
97 sPThrSerSerIleThrSerCysGlyAspGly 107
||||| ||||| ||||| ||||| |||||
2163 GACCTCTTCCACTTCTTCTACTTCTGCTGCTGCT 2132
```


This Page Blank (uspto)

1	leuGISeISeIySIyHISLySIySeIeueGleuIeueIuPnrySSeSe	17
1922	TtGGAAAGCAAGAGCTATAAAATTtCtGCAGTtACTtACtGtTtCTtC	1977
17	rAspAspAlngGlyHISerSeIeuthrAsuSeIProleuAspSerSeIc	34
1972	TtATtGCtCCGGGtCAtTtCTtCtTtGtACtCAAtCtCCtCCtTtAAtTtCAAtTt	2022
34	ylSIySIuSeISeISeIValSeIValtHrSeIProSeIglValSeISeISeI	50
2022	GtTAAAGAAAtTtCTtGtTtAGtGtACtCCAGtCCtCCtGtAGtGtCTtCCtCCtCT	2074

```
51 ThrSerGlyValSerSerThrSerAsnMetHisGlySerLeuLeu1 67
|||||
2072 ACATCTGGAGAGATACCTCTACATCCAAATATGATGGGTCACTGTTTACA 2121
67 nGluysHisArgIleLeuHisLysLeuLeuGlnAsnGlyAsnSerProA 84
|||||
2122 AGAAGAGCCCGGATTTTGCACAAGTTGCTGCAGAAATGGGAATTCACGAG 2171
84 IagIValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2172 CTGAGGTACCAAGATTACTGCACAAGCCACTGGGAAAGACACACAGCACT 2221
101 IleThrSerCysGlyAspGlyAsnValValLysGlnGluGlnLeuSerPr 117
|||||
2222 ATAACTCTTGTGGGACGGAATGTTGTCAAGCAGACAGCTAAGTCC 2271
117 oLysLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
|||||
2272 TAAGAAGAGGAGAAATAAATGCACTTCTTAGATACCTGCTGCACAGGATG 2321
134 sPProSerAspAlaLeuSer 140
|||||
2322 ATCTAGTATGCACTCTCT 2341

seq_name: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:US-09-513-066-1

seq_documentation_block:
: Sequence 1, Application US/09513066
: GENERAL INFORMATION:
: APPLICANT: Chen, J. Don
: APPLICANT: Leo, Christopher
: APPLICANT: Li, Hui
: TITLE OF INVENTION: NOVEL USES FOR THE RAC3 TRANSCRIPTIONAL COACTIVATOR OF
: TITLE OF INVENTION: STEROID NUCLEAR RECEPTORS
: FILE REFERENCE: UMG-026CP
: CURRENT APPLICATION NUMBER: US/09/513,066
: CURRENT FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: USSN 09/041,994
: PRIOR FILING DATE: 1998-03-13
: PRIOR APPLICATION NUMBER: USSN 60/073,674
: PRIOR FILING DATE: 1998-02-04
: NUMBER OF SEQ ID NOS: 37
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 4496
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (86)..(4330)
US-09-513-066-1

alignment_scores:
Quality: 703.00 Length: 140
Ratio: 5.021 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-041-994-2_COPY_613_752 x US-09-513-066-1 ..
Align seg 1/1 to: US-09-513-066-1 from: 1 to: 4496

1 LeuGluSerLysGlyHisLysLysLeuLeuGlnLeuLeuThrCysSerSe 17
|||||
1922 TTGGAAGAGCAAGGTCATATAAAATTACGCACTTACTTCTTCTTC 1971
17 TAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSerSerC 34
|||||
1972 TGATGACCGGGGTCAATCCCTTGACCAATCCGCCCTAGATTCAAGTT 2021
34 yLysGluSerSerValSerValThrSerProSerGlyValSerSerSer 50
```

```
|||||
2022 GTAAAGAAATCTTCTGTAGTGCACAGCCCTCTGAGACTCTCTCTCT 2071
51 ThrSerGlyValSerSerThrSerAsnMetHisGlySerLeuLeu1 67
|||||
2072 ACATCTGGAGAGATACCTCTACATCCAAATATGATGGGTCACTGTTTACA 2121
67 nGluysHisArgIleLeuHisLysLeuLeuGlnAsnGlyAsnSerProA 84
|||||
2122 AGAAGAGCCCGGATTTTGCACAAGTTGCTGCAGAAATGGGAATTCACGAG 2171
84 IagIValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2172 CTGAGGTACCAAGATTACTGCACAAGCCACTGGGAAAGACACACAGCACT 2221
101 IleThrSerCysGlyAspGlyAsnValValLysGlnGluGlnLeuSerPr 117
|||||
2222 ATAACTCTTGTGGGACGGAATGTTGTCAAGCAGACAGCTAAGTCC 2271
117 oLysLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
|||||
2272 TAAGAAGAGGAGAAATAAATGCACTTCTTAGATACCTGCTGCACAGGATG 2321
134 sPProSerAspAlaLeuSer 140
|||||
2322 ATCTAGTATGCACTCTCT 2341

seq_name: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:US-60-068-511-1

seq_documentation_block:
: Sequence 1, Application US/60068511
: GENERAL INFORMATION:
: APPLICANT: Suen, Chen-Shian
: APPLICANT: Freil, Donald E.
: APPLICANT: Lyttle, Richard C.
: TITLE OF INVENTION: Cloning and Expression of a Nuclear
: TITLE OF INVENTION: Receptor Coactivator Proteins and Uses Thereof
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: American Home Products Corporation
: STREET: One Campus Drive
: CITY: Parsippany
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07054
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/60/068,511
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Walsh, Andrea C.
: REGISTRATION NUMBER: 34,988
: REFERENCE/DOCKET NUMBER: 97243-00
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 973-683-2169
: TELEFAX: 973-683-4117
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4789 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: FEATURE:
```

NAME/KEY: CDS
LOCATION: 185..4750
US-60-068-511-1

alignment_scores:

Quality: 700.00 Length: 140
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.286

alignment_block:

US-09-041-994-2_COPY_613_752 x US-60-068-511-1

Align seg 1/1 to: US-60-068-511-1 from: 1 to: 4789

```

1 LeuGlSerLysGlyHisLysLysLeuGlnLeuLeuThrCysSerSe 17
|||||
2021 TTGGAAAGCAAGGTCATATAAAATTACTGCAGTACTTACTCTGTCCTTC
17 rAspAspArgLysHisSerSerLeuThrAsnSerProLeuAspSerSe 34
|||||
2071 TGATGACCGGGGTCATCTCTTGTGACCACTCCCTGAGATTCAGATT
34 ySLysGluSerSerValSerValThrSerProSerGlyValSerSerSe 50
|||||
2121 GTAAAGAAATCTTCTGTAGTGTCAACACACCCCTCGAGTCTCTCTCT
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeu 67
|||||
2171 ACATCTGGAGAGATCTCTTACATCCCAATATGATGGTGCATCTGTACA
67 nGluLysHisArgIleLeuHisLysLysLeuGlnAsnGlyAsnSerProA 84
|||||
2221 AGAAGACACCGGATTTGACAAAGTTGCTCAGAGATGGGAATTCACAG
84 laGluValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSe 100
|||||
2271 CTGAGGTAGCCAGATTTACTGCAGAGGCACTGGGAAAGACACAGCAGT
101 IleThrSerCysGlyAspGlyAsnValAlaLysGlnGlnLeuSerPr 117
|||||
2321 ATAACTCTTCTGTGGGAGCGGAAATGTTGTCAAGCAGAGCACTAAAGTC
117 oLysLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
|||||
2371 TAAGAAAGAGGAGATATATGCACTTCTTATGATACCTGCTGGACAGGAGT
134 sPProSerAspAlaLeuSer 140
|||||
2421 ATCCTAGTATGACACTCTCT 2440

```

seq_name: /cgn2_6/prodata/1/pna/US094_COMB.seq:us-09-440-612-1

seq_documentation_block:

```

; Sequence 1, Application US/09440612
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-3 EXPRESSION
; FILE REFERENCE: RTS-0042
; CURRENT APPLICATION NUMBER: US/09/440,612
; CURRENT FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 6754
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (184)..(4422)
US-09-440-612-1

```

alignment_scores:

Quality: 700.00 Length: 140
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.286

alignment_block:

US-09-041-994-2_COPY_613_752 x US-09-440-612-1

Align seg 1/1 to: US-09-440-612-1 from: 1 to: 6754

```

1 LeuGlSerLysGlyHisLysLysLeuGlnLeuLeuThrCysSerSe 17
|||||
2050 TTGGAAAGCAAGGTCATATAAAATTACTGCAGTACTTACTCTGTCCTTC
17 rAspAspArgLysHisSerSerLeuThrAsnSerProLeuAspSerSe 34
|||||
2100 TGATGACCGGGGTCATCTCTTGTGACCACTCCCTGAGATTCAGATT
34 ySLysGluSerSerValSerValThrSerProSerGlyValSerSerSe 50
|||||
2150 GTAAAGAAATCTTCTGTAGTGTCAACACACCCCTCGAGTCTCTCTCT
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeu 67
|||||
2200 ACATCTGGAGAGATATCTTACATCCCAATATGATGGTGCATCTGTACA
67 nGluLysHisArgIleLeuHisLysLysLeuGlnAsnGlyAsnSerProA 84
|||||
2250 AGAAGACACCGGATTTGACAAAGTTGCTCAGAGATGGGAATTCACAG
84 laGluValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSe 100
|||||
2300 CTGAGGTAGCCAGATTTACTGCAGAGGCACTGGGAAAGACACAGCAGT
101 IleThrSerCysGlyAspGlyAsnValAlaLysGlnGlnLeuSerPr 117
|||||
2350 ATAACTTCTTGTGGGAGCGGAAATGTTGTCAAGCAGAGCACTAAAGTC
117 oLysLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
|||||
2400 TAAGAAAGAGGAGATATATGCACTTCTTATGATACCTGCTGGACAGGAGT
134 sPProSerAspAlaLeuSer 140
|||||
2450 ATCCTAGTATGACACTCTCT 2469

```

seq_name: /cgn2_6/prodata/1/pna/US094_COMB.seq:us-09-440-612-3

seq_documentation_block:

```

; Sequence 3, Application US/09440612A
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-3 EXPRESSION
; FILE REFERENCE: RTS-0042
; CURRENT APPLICATION NUMBER: US/09/440,612A
; CURRENT FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 3
; LENGTH: 6754
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (184)..(4422)
US-09-440-612-3

```

alignment_scores:

Quality: 700.00 Length: 140
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.286

alignment_block:
US-09-041-994-2_COPY_613_752 x US-09-440-612-3 ..

Align seg 1/1 to: US-09-440-612-3 from: 1 to: 6754

```
1 LeuGluSerLysGlyHisLysLysLeuGlnLeuLeuThrCysSerse 17
|||||
2050 TTGGAAAGCAAGGCTATAAAATTACTGCAAGTACTTACTGTTCTTC 2099
17 rAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSerSerc 34
|||||
2100 TGATGACCGGGGTCATTCCTCTTGACCAACTCCCCCTAGATTCAAGTT 2149
34 yslYsluSerSerValSerValThrSerProSerGlyValSerSerSer 50
|||||
2150 GTAAAGAAATCTTCTGTAGTGTACACAGCCCCCTGTGAGTCTCTCTCT 2199
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeuG 67
|||||
2200 ACATCTGGAGGAGTATCCTTACATCCAAATATGCAATGGGTCACCTGTTACA 2249
67 ngLuLySHisArgIleLeuHisLysLeuLeuGlnAsnGlyAsnSerPro 84
|||||
2250 AGAAGAACACCGGATTTTGACAAAGTTGTCGCAAGATGGGAATTCACCA 2299
84 laGluValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2300 CTGAGGTATCCAGCAAGATTACTGCAAGAGCCACTGGAAAGACACACAGCT 2349
101 IleThrSerCysGlyAspGlyAsnValAlaLysGlnGluGlnLeuSerP 117
|||||
2350 ATTAACCTCTTGTGTGGGACGGAATGTTGTCAACAGCAGACGACTAAGTCC 2399
117 oLySLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAsp 134
|||||
2400 TAAGAAGAAAGCAAGATTAATGCACTTCTTAGATACCTGCTGACAGGAGATG 2449
134 sPProSerAspAlaLeuSer 140
|||||
2450 ATCTAGTGTGTCACCTCTCT 2469
```

seq_name: /cgn2_6/ptodata/1/pna/US091_COMB.seq:US-09-125-635-1

seq_documentation_block:
; Sequence 1, Application US/09125635
; GENERAL INFORMATION:
; APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE
; TITLE OF INVENTION: AlB1, A novel steroid receptor co-activator
; FILE REFERENCE: 49944
; CURRENT APPLICATION NUMBER: US/09/125_635
; CURRENT FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/049_728
; PRIOR FILING DATE: 1997-06-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 6835
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(4463)
US-09-125-635-1

alignment_scores:
Quality: 700.00 Length: 140
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.286

alignment_block:
US-09-041-994-2_COPY_613_752 x US-09-125-635-1 ..

Align seg 1/1 to: US-09-125-635-1 from: 1 to: 6835

```
1 LeuGluSerLysGlyHisLysLysLeuGlnLeuLeuThrCysSerse 17
|||||
2037 TTGGAAAGCAAGGCTATAAAATTACTGCAAGTACTTACTGTTCTTC 2086
17 rAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSerSerc 34
|||||
2087 TGATGACCGGGGTCATTCCTCTTGACCAACTCCCCCTGATTTCAAGTT 2136
34 yslYsluSerSerValSerValThrSerProSerGlyValSerSerSer 50
|||||
2137 GTAAAGAAATCTTCTGTAGTGTACCAAGCCCCCTGTGAGTCTCTCTCT 2186
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeuG 67
|||||
2187 ACATCTGGAGGAGTATCCTTACATCCAAATATGCAATGGGTCACCTGTTACA 2236
67 ngLuLySHisArgIleLeuHisLysLeuLeuGlnAsnGlyAsnSerPro 84
|||||
2237 AGAAGAACACCGGATTTTGACAAAGTTGTCGCAAGATGGGAATTCACCA 2286
84 laGluValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2287 CTGAGGTATCCAGCAAGATTACTGCAAGAGCCACTGGAAAGACACACAGCT 2336
101 IleThrSerCysGlyAspGlyAsnValAlaLysGlnGluGlnLeuSerP 117
|||||
2337 ATTAACCTCTTGTGTGGGACGGAATGTTGTCAACAGCAGACGACTAAGTCC 2386
117 oLySLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAsp 134
|||||
2387 TAAGAAGAAAGCAAGATTAATGCACTTCTTAGATACCTGCTGACAGGAGATG 2436
134 sPProSerAspAlaLeuSer 140
|||||
2437 ATCTAGTGTGTCACCTCTCT 2456
```

seq_name: /cgn2_6/ptodata/1/pna/US093_COMB.seq:US-09-359-922-3962

seq_documentation_block:
; Sequence 3962, Application US/09359922
; GENERAL INFORMATION:
; APPLICANT: HysSeq, Inc.
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-752CON1
; CURRENT APPLICATION NUMBER: US/09/359_922
; CURRENT FILING DATE: 1999-07-22
; EARLIER APPLICATION NUMBER: US 09/205_155
; EARLIER FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 13203
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3962
; LENGTH: 6855
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(6855)
; OTHER INFORMATION: n = A,T,C or G
US-09-359-922-3962

alignment_scores:
Quality: 700.00 Length: 140
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.286

alignment_block:
US-09-041-994-2_COPY_613_752 x US-09-359-922-3962 ..

Align seg 1/1 to: US-09-359-922-3962 from: 1 to: 6855


```

1 LeuGlusertLysGlyHisLysLysLeuLeuGlnLeuLeuThrCysSerSe 17
  |||||||
2037 TTGAAAGCAAAAGGTCATATAAAAAATTACTGCAGATTACTTACCTGTTCTTC 2086
17 RASPAAPRGlyHisSerSerLeuThrAsnSerProLeuAspSerSerc 34
  |||||||
2087 TGATGACCGGGGTCATTCCTTGCACCAACTCCCCCTAGATTCAAGTT 2136
34 ySLysGlySerSerValSerValThrSerProSerGlyValSerSerSer 50
  |||||||
2137 GTAAGAAATCTTCTGTAGTGTACACAGCCCTCTGGAGTCTCCTCTCT 2186
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeuG 67
  |||||||
2187 ACATCTGGAGAGATACCTCTACATCCAAATATGCATGGGTCACTGTATCA 2236
67 nGluLysHisArgIleLeuHisLysLeuLeuGlnAsnGlyAsnSerProA 84
  |||||||
2237 AGAAGACACCGGATTTTGCACAAAGTTCTGCAGAAATGGGAATTCACCA 2286
84 IagIuValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
  |||||||
2287 CTGAGGTAGCCCAAGATTACTGCAGAACCCACTGGGAAAGACACACAGCA 2336
101 IleThrSerCysGlyAspGlyAsnValValLysGlnGlnLeuSerPr 117
  |||||||
2337 ATAACCTTCTGTGGGACGAAATGTTGTCAAGACGAGCAGCTAAAGTCC 2386
117 OlySLysLysGlyLysAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
  |||||||
2387 TAAAGAGAGAGAGATATATGCACTTCTTAGATACTGCTGACACAGGATG 2436
134 sPProSerAspAlaLeuSer 140
  |||||||
2437 ATCCTAGTAGATGCACTCTCT 2456

```

seq_name: /cgn2_6/ptodata/1/pna/US093_COMB.seq:us-09-359-922-3962

```

seq_documentation_block:
: Sequence 3962, Application US/09359922A
: GENERAL INFORMATION:
: APPLICANT: Ieshkowitz, Dena
: APPLICANT: Liu, Jin
: TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
: FILE REFERENCE: 20411-752CON1
: CURRENT APPLICATION NUMBER: US/09/359,922A
: EARLIER FILING DATE: 1999-07-22
: EARLIER APPLICATION NUMBER: US 09/205,155
: EARLIER FILING DATE: 1998-12-03
: EARLIER APPLICATION NUMBER: US 09/034,341
: NUMBER OF SEQ ID NOS: 13203
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 3962
: LENGTH: 6855
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)-(6855)
: OTHER INFORMATION: n = A,T,C or G
US-09-359-922-3962

```

alignment_scores: Quality: 700.00 Length: 140
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.286

alignment_block: US-09-041-994-2_copy_613_752 x US-09-359-922-3962

```

Align seg 1/1 to: US-09-359-922-3962 from: 1 to: 6855
1 LeuGlusertLysGlyHisLysLysLeuLeuGlnLeuLeuThrCysSerSe 17
  |||||||
2037 TTGAAAGCAAAAGGTCATATAAAAAATTACTGCAGATTACTTACCTGTTCTTC 2086
17 RASPAAPRGlyHisSerSerLeuThrAsnSerProLeuAspSerSerc 34
  |||||||
2087 TGATGACCGGGGTCATTCCTTGCACCAACTCCCCCTAGATTCAAGTT 2136
34 ySLysGlySerSerValSerValThrSerProSerGlyValSerSerSer 50
  |||||||
2137 GTAAGAAATCTTCTGTAGTGTACACAGCCCTCTGGAGTCTCCTCTCT 2186
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeuG 67
  |||||||
2187 ACATCTGGAGAGATACCTCTACATCCAAATATGCATGGGTCACTGTATCA 2236
67 nGluLysHisArgIleLeuHisLysLeuLeuGlnAsnGlyAsnSerProA 84
  |||||||
2237 AGAAGACACCGGATTTTGCACAAAGTTCTGCAGAAATGGGAATTCACCA 2286
84 IagIuValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
  |||||||
2287 CTGAGGTAGCCCAAGATTACTGCAGAACCCACTGGGAAAGACACACAGCA 2336
101 IleThrSerCysGlyAspGlyAsnValValLysGlnGlnLeuSerPr 117
  |||||||
2337 ATAACCTTCTGTGGGACGAAATGTTGTCAAGACGAGCAGCTAAAGTCC 2386
117 OlySLysLysGlyLysAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
  |||||||
2387 TAAAGAGAGAGAGATATATGCACTTCTTAGATACTGCTGACACAGGATG 2436
134 sPProSerAspAlaLeuSer 140
  |||||||
2437 ATCCTAGTAGATGCACTCTCT 2456

```

seq_name: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:US-60-209-009-198

```

seq_documentation_block:
: Sequence 198, Application US/60209009
: GENERAL INFORMATION:
: APPLICANT: Faris, Mary
: APPLICANT: Pearson, Cecelia I.
: TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
: FILE REFERENCE: PA-0027 P
: CURRENT APPLICATION NUMBER: US/60/209,009
: EARLIER FILING DATE: 2000-06-01
: NUMBER OF SEQ ID NOS: 501
: SOFTWARE: PERL Program
: SEQ ID NO 198
: LENGTH: 7116
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 1094199.1
: NAME/KEY: unsure
: LOCATION: 3941-3993, 5899-5939, 6951
: OTHER INFORMATION: a, t, c, g, or other
US-60-209-009-198

```

alignment_scores: Quality: 700.00 Length: 140
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.286

alignment_block: US-09-041-994-2_copy_613_752 x US-60-209-009-198

Align seg 1/1 to: US-60-209-009-198 from: 1 to: 7116

```
1 LeuGluSerLysGlyHisLysLysLeuGlnLeuLeuThrCysSerse 17
|||||
2037 TTGGAAAGCAAAAGGTCATAAAAAATTACTGACGTTACTTACCTGTTCTTCT 2086
17 rAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSerSerC 34
|||||
2087 TGATGACCGGGGTCATTCTCTTGACCAATCTCCCTAGATTCAAGTT 2136
34 yLysGluSerSerValSerValThrSerProSerGlyValSerSerSer 50
|||||
2137 GTAAGAAATCTTCTGTAGTCACACAGCCCTCTGAGTCTCTCTCTCT 2186
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuGln 67
|||||
2187 ACATCTGGAGGAGTATCTCTACATCCAAATATGCAATGGCTCACTGTACA 2236
67 nGluLysHisArgGlyLeuHisLysLeuGlnAsnGlyAsnSerProA 84
|||||
2237 AGAGAAAGCACCGGATTTCACAAAGTTGCTGCAGAAATGGAAATTCACAG 2286
84 lAgLysValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2287 CTGAGGTACCAAGATTACTGCAGAGCCACTGGAAAGACACACACAGCT 2336
101 lIleThrSerCysGlyAspGlyAsnValValLysGlnGlnLeuSerPr 117
|||||
2337 ATAACTCTTGTGGGCGACGAAATGTTGTCAGACAGCAGCAGCTAAGTTC 2386
117 oLysLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
|||||
2387 TAAGAAAGAGAGAAATATGCACTTCTTAGATACCTGCTGACAGGAGATG 2436
134 sPProSerAspAlaLeuSer 140
|||||
2437 ATCCTAGTGTATGCATCTCT 2456
```

seq_name: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:US-60-213-360-7991

```
seq_documentation_block:
: Sequence 7991, Application US/60213360
: GENERAL INFORMATION:
: APPLICANT: Morris, MacDonald
: APPLICANT: Lal, Preclt
: TITLE OF INVENTION: Method for the identification of sequence polymorphisms using
: TITLE OF INVENTION: Polynucleotide sequence databases, and single nucleotide polymor
: FILE REFERENCE: GX-0014 P
: CURRENT APPLICATION NUMBER: US/60/213, 360
: CURRENT FILING DATE: 2000-06-21
: NUMBER OF SEQ ID NOS: 8347
: SOFTWARE: PERL Program
: SEQ ID NO 7991
: LENGTH: 7116
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Incyte ID No: 1094199.1
: NAME/KEY: unsure
: LOCATION: 3941-3993, 5899-5939, 6951
: OTHER INFORMATION: a, t, c, g, or other
US-60-213-360-7991
```

alignment_scores:

Quality: 700.00 Length: 140
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.286

alignment_block:

US-09-041-994-2_COPY_613_752 x US-60-213-360-7991 ..

Align seg 1/1 to: US-60-213-360-7991 from: 1 to: 7116

```
1 LeuGluSerLysGlyHisLysLysLeuGlnLeuLeuThrCysSerse 17
|||||
2037 TTGGAAAGCAAAAGGTCATAAAAAATTACTGACGTTACTTACCTGTTCTTCT 2086
17 rAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSerSerC 34
|||||
2087 TGATGACCGGGGTCATTCTCTTGACCAATCTCCCTAGATTCAAGTT 2136
34 yLysGluSerSerValSerValThrSerProSerGlyValSerSerSer 50
|||||
2137 GTAAGAAATCTTCTGTAGTCACACAGCCCTCTGAGTCTCTCTCTCTCT 2186
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuGln 67
|||||
2187 ACATCTGGAGGAGTATCTCTACATCCAAATATGCAATGGCTCACTGTACA 2236
67 nGluLysHisArgGlyLeuHisLysLeuGlnAsnGlyAsnSerProA 84
|||||
2237 AGAGAAAGCACCGGATTTCACAAAGTTGCTGCAGAAATGGAAATTCACAG 2286
84 lAgLysValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2287 CTGAGGTACCAAGATTACTGCAGAGCCACTGGAAAGACACACACAGCT 2336
101 lIleThrSerCysGlyAspGlyAsnValValLysGlnGlnLeuSerPr 117
|||||
2337 ATAACTCTTGTGGGCGACGAAATGTTGTCAGACAGCAGCAGCTAAGTTC 2386
117 oLysLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
|||||
2387 TAAGAAAGAGAGAAATATGCACTTCTTAGATACCTGCTGACAGGAGATG 2436
134 sPProSerAspAlaLeuSer 140
|||||
2437 ATCCTAGTGTATGCATCTCT 2456
```

seq_name: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:US-09-652-123-8947

```
seq_documentation_block:
: Sequence 8947, Application US/09652123
: GENERAL INFORMATION:
: APPLICANT: Holtzman, Douglas A.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
: TITLE OF INVENTION: THEREFOR
: FILE REFERENCE: 1600,1186-001
: CURRENT APPLICATION NUMBER: US/09/652,123
: CURRENT FILING DATE: 2000-08-30
: PRIOR APPLICATION NUMBER: 60/151,135
: PRIOR FILING DATE: 1999-08-30
: NUMBER OF SEQ ID NOS: 9796
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 8947
: LENGTH: 7888
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-652-123-8947
```

alignment_scores:

Quality: 700.00 Length: 140
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.286

alignment_block:

US-09-041-994-2_COPY_613_752 x US-09-652-123-8947 ..

Align seg 1/1 to: US-09-652-123-8947 from: 1 to: 7888

1 LeuGluSerLysGlyHisLysLysLeuGlnLeuLeuThrCysSerse 17

```
2048 TTGAAAGCAAGCTCATTAATAAATTAAGTACAGTACTACTTCTCTC 2097
17 rAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSerSerC 34
2148 GTAAAGATCTTCTGTAGTGTACACAGCCCTCTGGAGTCTCTCTCT 2197
2098 TGATGAGCGGGGTCATCTCTTCTGACCACTCCCTTACATTCAGATT 2147
34 yslYsgIuSerSerValSerValThrSerProSerGlyValSerSerSer 50
2148 GTAAAGATCTTCTGTAGTGTACACAGCCCTCTGGAGTCTCTCTCT 2197
51 ThrSerGlyValSerSerThrSerAsnMetHisGlySerLeuLeuG 67
2198 ACATCTGAGAGAGATACCTCTACATCCAAATTCAGTGTACCTTACA 2247
67 nGUlySHisArgIleLeuHisLysLeuLeuGlnAsnGlyAsnSerProA 84
2248 AGAAGACACCGGATTTTGCACAAAGTGTGCAAGATGGGAATTCACCA 2297
84 laGIuValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
2298 CTGAGTAGCCAGATTAATCTGACAGACCCCTGGGAAAGACACACAGAT 2347
101 IleThrSerCysGlyAspGlyAsnValValLysGlnGluGlnLeuSerPr 117
2348 ATAACTTCTTGTGGGAGCGAAATGTTGTCAAGCAGAGCAGCTAAAGTCC 2397
117 OlySLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
2398 TAAGAAAGAGAGAAATATATGACTTCTTACATCTCTGACAGGATG 2447
134 sPProSerAspAlaLeuSer 140
2448 ATCTAGTATGACACTCTCT 2467
```

seq_name: /cgn2_6/plodata/1/pna/US096C_COMB.seq:US-09-652-127-9831

```
seq_documentation_block:
; Sequence 9831, Application US/09652127
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1183-001
; CURRENT APPLICATION NUMBER: US/09/652,127
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,134
; NUMBER OF SEQ ID NOS: 10475
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9831
; LENGTH: 7888
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-127-9831
```

alignment_scores:
Quality: 700.00 Length: 140
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.286

alignment_block:
US-09-041-994-2_copy_613_752 x US-09-652-127-9831

Align seg 1/1 to: US-09-652-127-9831 from: 1 to: 7888

```
1 LeuGluSerLysGlyHisLysLysLeuLeuGlnLeuLeuThrCysSerSe 17
2048 TTGAAAGCAAGCTCATTAATAAATTAAGTACAGTACTACTTCTCTC 2097
17 rAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSerSerC 34
```

```
2098 TGATGACCGGGGTCATCTCTCTTGTACCACTCCCTTACATTCAGATT 2147
34 yslYsgIuSerSerValSerValThrSerProSerGlyValSerSerSer 50
2148 GTAAAGATCTTCTGTAGTGTACACAGCCCTCTGGAGTCTCTCTCTCT 2197
51 ThrSerGlyValSerSerThrSerAsnMetHisGlySerLeuLeuG 67
2198 ACATCTGAGAGAGATACCTCTACATCCAAATTCAGTGTACCTTACA 2247
67 nGUlySHisArgIleLeuHisLysLeuLeuGlnAsnGlyAsnSerProA 84
2248 AGAAGACACCGGATTTTGCACAAAGTGTGCAAGATGGGAATTCACCA 2297
84 laGIuValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
2298 CTGAGTAGCCAGATTAATCTGACAGACCCCTGGGAAAGACACACAGAT 2347
101 IleThrSerCysGlyAspGlyAsnValValLysGlnGluGlnLeuSerPr 117
2348 ATAACTTCTTGTGGGAGCGAAATGTTGTCAAGCAGAGCAGCTAAAGTCC 2397
117 OlySLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
2398 TAAGAAAGAGAGAAATATATGACTTCTTACATCTCTGACAGGATG 2447
134 sPProSerAspAlaLeuSer 140
2448 ATCTAGTATGACACTCTCT 2467
```

seq_name: /cgn2_6/plodata/1/pna/US096C_COMB.seq:US-09-652-355-11058

```
seq_documentation_block:
; Sequence 11058, Application US/09652355
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1192-001
; CURRENT APPLICATION NUMBER: US/09/652,355
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,136
; NUMBER OF SEQ ID NOS: 11227
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11058
; LENGTH: 7888
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-355-11058
```

alignment_scores:
Quality: 700.00 Length: 140
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.286

alignment_block:
US-09-041-994-2_copy_613_752 x US-09-652-355-11058

Align seg 1/1 to: US-09-652-355-11058 from: 1 to: 7888

```
1 LeuGluSerLysGlyHisLysLysLeuLeuGlnLeuLeuThrCysSerSe 17
2048 TTGAAAGCAAGCTCATTAATAAATTAAGTACAGTACTACTTCTCTC 2097
17 rAspAspArgGlyHisSerSerLeuThrAsnSerProLeuAspSerSerC 34
2098 TGATGACCGGGGTCATCTCTTGTGACCACTCCCTTACATTCAGATT 2147
34 yslYsgIuSerSerValSerValThrSerProSerGlyValSerSerSer 50
2148 GTAAAGATCTTCTGTAGTGTACACAGCCCTCTGGAGTCTCTCTCTCT 2197
```

```
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeu1 67
|||||
2198 ACATCTGGAGGAGTATCTCTACATCCAAATATGCAATGGGTCACTGTACA 2247
67 nCluYsNHsArgIleuNHsIleuNHsLeuNHsngIYsAnSerProA 84
|||||
2248 AGAGAAGCACCGGATTTTGCACAAAGTTGCTGCAGAAATGGGAATTCACGAG 2297
84 laGlValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2298 CTGAGGTACCAAGATTACTGCAGAAAGCCACTGGGAAAGACACAGCACT 2347
101 IleThrSerCysGlyAspGlyAsnValValLysGlnGlnLeuSerPr 117
|||||
2348 ATAACTTCTTGTGGGACGGAATGTTGTCAAGCAGAGCACGACTAAGTCC 2397
117 oLysIlySlySguAsnAsnAlaLeuNHsArgTyIreLeuAspArgAspA 134
|||||
2398 TAAGAAAGAGGAGAAATATGCACTTCTTAGATACCTGCTGCAGAGGAGT 2447
134 sPProSerAspAlaLeuSer 140
|||||
2448 ATCTAGTGATGCACTCTCT 2467
```

seq_name: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:US-09-663-693-1028

```
seq_documentation_block:
: Sequence 1028, Application US/09663693
: GENERAL INFORMATION:
: APPLICANT: Gearing, David P.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
: FILE OF INVENTION: THEREFOR
: FILE REFERENCE: 1600.1195-001
: CURRENT APPLICATION NUMBER: US/09/663,693
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: 60/154,986
: PRIOR FILING DATE: 1999-09-21
: NUMBER OF SEQ ID NOS: 1340
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1028
: LENGTH: 7888
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-663-693-1028
```

alignment_scores:
Quality: 700.00 Length: 140
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.286

alignment_block:
US-09-041-994-2_COPY_613_752 x US-09-663-693-1028 ..

Align seg 1/1 to: US-09-663-693-1028 from: 1 to: 7888

```
1 LeuGluSerIleGlyNHsIlySlyLeuNHsGlnLeuNHsThrCysSerSe 17
|||||
2048 TTGGAAGACAAAGGTCATATAAAATTAAGTCACTTACTTACTGTTCTTC 2097
17 rAspAspArgGlyNHsSerSerLeuNHsThrAsnSerProLeuAspSerSe 34
|||||
2098 TGATGACCGGGTCATCTCTCTTGACCAACATCCCCCTAGATTCAAGTT 2147
34 yAlSlyGluSerSerValSerValIThrSerProSerGlyValSerSerSe 50
|||||
2148 GTAAAGAAATCTTCTGTAGTGTACACAGCCCTCTGGAATCTCTCTCT 2197
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeu1 67
|||||
2198 ACATCTGGAGGAGTATCTCTACATCCAAATATGCAATGGGTCACTGTACA 2247
```

```
67 nCluYsNHsArgIleuNHsIleuNHsLeuNHsngIYsAnSerProA 84
|||||
2248 AGAGAAGCACCGGATTTTGCACAAAGTTGCTGCAGAAATGGGAATTCACGAG 2297
84 laGlValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2298 CTGAGGTACCAAGATTACTGCAGAAAGCCACTGGGAAAGACACAGCACT 2347
101 IleThrSerCysGlyAspGlyAsnValValLysGlnGlnLeuSerPr 117
|||||
2348 ATAACTTCTTGTGGGACGGAATGTTGTCAAGCAGAGCACGACTAAGTCC 2397
117 oLysIlySlySguAsnAsnAlaLeuNHsArgTyIreLeuAspArgAspA 134
|||||
2398 TAAGAAAGAGGAGAAATATGCACTTCTTAGATACCTGCTGCAGAGGAGT 2447
134 sPProSerAspAlaLeuSer 140
|||||
2448 ATCTAGTGATGCACTCTCT 2467
```

seq_name: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:US-09-698-014-5135

```
seq_documentation_block:
: Sequence 5135, Application US/09698014
: GENERAL INFORMATION:
: APPLICANT: Gearing, David P.
: APPLICANT: Holtzman, Douglas A.
: APPLICANT: Kingsbury, Gillian A.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
: FILE OF INVENTION: THEREFOR
: FILE REFERENCE: 1600.2014-001
: CURRENT APPLICATION NUMBER: US/09/698,014
: PRIOR FILING DATE: 2000-10-27
: PRIOR APPLICATION NUMBER: 60/162,363
: PRIOR FILING DATE: 1999-10-29
: NUMBER OF SEQ ID NOS: 6098
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5135
: LENGTH: 7888
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-698-014-5135
```

alignment_scores:
Quality: 700.00 Length: 140
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.286

alignment_block:
US-09-041-994-2_COPY_613_752 x US-09-698-014-5135 ..

Align seg 1/1 to: US-09-698-014-5135 from: 1 to: 7888

```
1 LeuGluSerIleGlyNHsIlySlyLeuNHsGlnLeuNHsThrCysSerSe 17
|||||
2048 TTGGAAGACAAAGGTCATATAAAATTAAGTCACTTACTTACTGTTCTTC 2097
17 rAspAspArgGlyNHsSerSerLeuNHsThrAsnSerProLeuAspSerSe 34
|||||
2098 TGATGACCGGGTCATCTCTCTTGACCAACATCCCCCTAGATTCAAGTT 2147
34 yAlSlyGluSerSerValSerValIThrSerProSerGlyValSerSerSe 50
|||||
2148 GTAAAGAAATCTTCTGTAGTGTACACAGCCCTCTGGAATCTCTCTCT 2197
51 ThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeu1 67
|||||
2198 ACATCTGGAGGAGTATCTCTACATCCAAATATGCAATGGGTCACTGTACA 2247
67 nCluYsNHsArgIleuNHsIleuNHsLeuNHsngIYsAnSerProA 84
|||||
2248 AGAGAAGCACCGGATTTTGCACAAAGTTGCTGCAGAAATGGGAATTCACGAG 2297
```

```
84  IagluValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSer 100
|||||
2298 CTGAGGTAGCCAGATTACTGCAGAACCCACTGGGAAGACACACGACGT 2347
|||||
101  IleThrSerCysGlyAspGlyAsnValValLysGlnGlnLeuSerPr 117
|||||
2348 ATAACTTCTTGTTGGGACGGAAATGTTGTCACGACGACGAGCTAAAGTCC 2397
|||||
117  OlyLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspA 134
|||||
2398 TAAGAAAGAAAGAGATATATGCACTTCTTAGATACCTGCGACACGCGATG 2447
|||||
134  sPProSerAspAlaLeuSer 140
|||||
2448 ATCCTAGTGATGCACCTCT 2467
```

This Page Blank (uspto)

OM of: US-09-041-994-2_COPY_613_752 to: Pending_Patents_NA_New:* out_format : pfs
Date: May 1, 2001 9:18 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-196-296B-1 + 700.00 1304.56 3.0e-65 47
-DB=Pending_Patents_NA_New -OEM=fastap -SUFFIX=p2n.rnpn
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000
-LOOPEXT=0.000 -OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000
-XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1
-MATRIX=blomsun62 -TRANS=human40.cdi -LIST=45 -DOCALLGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=pfs -NORM=ext -MINLEN=0 -MAXLEN=200000000
-USER=US09041994_@cgn1_1_175 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPHY
-WAIT -THREADS=1

Search information block:
Query: US-09-041-994-2_COPY_613_752
Query length: 140
Database: Pending_Patents_NA_New:*
Database sequences: 201016
Database length: 146721182
Search time (sec): 114.180000

score list:

Sequence	Strid Orig	ZScore	Score Len	Documentation
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-196-296B-1 +		700.00	1304.56	3.0e-65 47
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12207 -		99.00	102.17	280.06
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12210 -		87.50	83.52	3.1e+03
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12210 -		86.00	80.66	4.4e+03
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12210 -		85.50	131.12	6.83
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12214 -		85.50	84.48	2.7e+03
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12212 -		85.00	87.55	1.8e+03
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12218 +		83.00	74.93	9.1e+03
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12212 +		82.00	81.82	3.8e+03
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12208 +		80.00	75.42	4.9e+03
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-08-956-171C-480 +		78.50	116.26	45.93
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-08-956-171C-420 +		78.50	109.22	113.11
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-739-449-218 +		78.00	72.84	1.2e+04
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12216 -		78.00	67.00	2.5e+04
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-08-956-171C-2968 +		77.00	133.33	5.14
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-08-956-171C-10 +		77.00	132.22	5.93
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-813-206-736 +		77.00	124.13	16.75
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-06-281-593-108 +		76.50	117.56	38.88
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12213 -		76.50	66.22	2.7e+04
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12215 +		76.00	61.69	4.7e+04
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-540-212A-32765 +		75.50	141.43	1.82
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-801-833-6569 +		75.50	124.07	16.88
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-811-380-681 +		75.00	116.43	44.95
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12211 +		74.50	65.13	3.1e+04
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12207 +		74.50	55.43	9.8e+04
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-06-248-505-604 +		74.00	88.58	1.6e+03
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-801-833-8092 -		73.50	117.13	41.08
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-801-833-8092 -		73.50	60.83	2.0e+04
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12216 +		73.50	68.01	5.8e+04
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12217 +		73.50	56.92	8.3e+04
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-08-956-171C-4457 +		73.00	132.49	5.73
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-423-844-169 +		73.00	114.83	55.22
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-423-844-169 +		73.00	95.79	634.63
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-06-248-823-11 -		73.00	78.76	5.6e+03
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12204 -		73.00	70.93	1.5e+04
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-739-449-211 +		72.50	69.72	1.8e+04
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-033-12204 -		72.00	113.87	62.41
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-08-956-171C-825 -		72.00	120.79	25.70

/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-663-561-6 +	72.00	105.15	191.08
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-663-561-32 +	72.00	103.27	243.12
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-663-561-33 +	72.00	103.07	249.49
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-08-956-171C-63 +	72.00	101.38	309.95
/cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-08-956-171C-63 +	72.00	93.88	810.54

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-196-296B-1

seq_documentation_block:
; Sequence 1, Application US/09196296B
; GENERAL INFORMATION:
; APPLICANT: Suen, Chen-Shian
; APPLICANT: Frai, Donald E.
; APPLICANT: Lytle, Richard C.
; TITLE OF INVENTION: Cloning and Expression of a Nuclear
; FILE REFERENCE: 0630/01376
; CURRENT APPLICATION NUMBER: US/09/196, 296B
; CURRENT FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4789
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-196-296B-1

alignment_scores:

Quality:	700.00	Length:	140
Ratio:	5.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	99.286

alignment_block:
US-09-041-994-2_COPY_613_752 x US-09-196-296B-1 ...

Align seg 1/1 to: US-09-196-296B-1 from: 1 to: 4789

1	leuGlSerLysGlyHisLysLysLeuLeuLeuLeuThrCysSer	17
2021	TTGGAAAGCAAGGCTATTAATAATTAAGTCTTACTTCTCTTC	2070
17	raspAargLysHisSerLeuThrAsnSerProLeuAspSerC	34
2071	TGAGACCGGGGCTATTCCTTGCACCACTCCCTTCAATTCAGTT	2120
34	ylsGlSerSerValSerValThrSerProSerGlyValSerSer	50
2121	GTAAGATCTTCTGTAAGTCAACGACCCCTTGAGTCTCTCTCT	2170
51	ThrSerGlyLysLysSerThrSerAsnMetHisLysSerLeuLeu	67
2171	ACATCTGAGGAGATGATCTTACATCCAAATGATGATGATGATG	2220
67	ngLysHisArgLysLeuHisLysLeuLeuLeuGlnAsnGlySerPro	84
2221	AGAAAGACCGGATTTGCAAGATGCTGCAAAATGGGATTCACCG	2270
84	lactValAlaLysLysLeuThrAlaGlnAlaThrGlyLysAspThrSer	100
2271	CTGAGTAGGACCAATTAATTAATTAATTAATTAATTAATTAAT	2320
101	lIleThrSerCysGlyAspGlyAsnValLysLysLysLysLysSer	117
2321	ATAACTCTTCTGAGGAGGAAATGTTCTCAAGCAGGAGGCTAATC	2370
117	oLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys	134
2371	TAAAGAGAGGAGATTAATTAATTAATTAATTAATTAATTAAT	2420
134	spproSerAspAlaLeuSer	140
2421	ATCTAGTAGTACACTCTCT	2440

```
seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12207
seq_documentation_block:
; Sequence 12207, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12207
; LENGTH: 1531974
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12207

alignment_scores:
Quality: 99.00 Length: 120
Ratio: 1.414 Gaps: 4
Percent Similarity: 58.333 Percent Identity: 26.667

alignment_block:
US-09-041-994-2_COPY_613_752 x US-09-335-032-12207/rev ..
Align seg 1/1 to reverse of: US-09-335-032-12207 from: 1 to: 1531974

21 GYHHSerSerLeuThrAsnSerProLeuAspSerSerCysLysGluSe 37
|||||
59821 GGTGATTCAGAGCTACACCTCCCTGTCACCTACACACATGATGCTATGC 59772
37 rSerValSerValThrSerProSerGlyVal.....SerSerSerThs 52
|||||
59771 CAGTGTGAAGAACACATCTCTCCATACCATATCATATGACATTCTAAAGCA 59722
52 erGlyGlyValSerSerThrSerAsnMetHisGlySerLeu...LeuGln 67
::|||
59721 ATACTGGTCCATCAGCAGCTGACATGCTGTTCAAGCAGCTGCTCTAA 59672
68 GILYSHISArgIleLeuHisLysLeuGlnAsnGlyAsnSerProAl 84
|||
59671 ACTACCACCCGTTAGCCATGTCATCTGGTCAAGCAGCTGCTCTAA 59622
84 aGluValAlaLysIleThrAlaGlnAlaThrGlyLysAspThrSerSerI 101
|||||
59621 ATCCACCGCTGTGCACAGGATCATTCAGTACTAAACTAACCTCTCAG 59572
101 leThrSerCysGlyAspGlyAsnValIleLysGlnGluGlnLeuSerPro 117
::|||
59571 TTACTCAT...GGACATACCACTGTT..... 59549
118 LysLysLysGluAsnAsnAlaLeuLeuArgTyrLeuLeuAspArgAspAs 134
|||||
59548 .....AAGATTAACCTCTCAGCTACGAGGCTTATTTCGACACATGATTC 59505
134 pProSerAsp 137
::|||
59504 CAACCTCAGAT 59495

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12210
seq_documentation_block:
; Sequence 12210, Application US/09335032
```

```
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12210
; LENGTH: 1090936
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12210

alignment_scores:
Quality: 87.50 Length: 166
Ratio: 0.941 Gaps: 10
Percent Similarity: 56.024 Percent Identity: 27.711

alignment_block:
US-09-041-994-2_COPY_613_752 x US-09-335-032-12210 ..
Align seg 1/1 to: US-09-335-032-12210 from: 1 to: 1090936

16 SerSerAspArgGlyHisSerSerLeuThrAsnSerPro.....Le 30
::|||
16167 AACAGTAATCTGTCAGAGCTCATCATCAATGAGCAATATCCGCGAGCT 161216
30 uAspSerSerCysLysGluSerSerValSerValThrSerProSerGly 47
|::|
161217 CAGTATGCGCTCAAGAACGCAATATA...TTAACTGCGCCAGCGCTTA 161263
47 aLSerSerSerThrSerGlyGlyValSerSerThSerAsn.....Met 61
::|||
161264 AAACAGACCATACACACAGCCGTAGATAAAGCTTCAATATATATAC 161313
62 HisGlySerLeuLeuGlnGluLys.....HisArgIleLe 73
|::|
161314 CATTAATAGCGCTTACATAGAGCTTGAAGAAGATATGTGTGCT 161363
73 uHisLysLeuLeuGlnAsnGlyAsnSerProAla...GluValAlaLysI 89
::|||
161364 AGAGAAGAAATCGGTTGAATTCATTGCGCAGCAGGTTGCTCA 161412
89 leThrAlaGlnAlaThrGlyLysAspThrSerSerIle.....Thr 102
::|||
161413 .....GCAGATTCAATCCAAATCTATTATAGCACCCCACT 161448
103 SerCysGlyAspGlyAsnValIleLysGlnGlnLeuSerProLysLys 119
|::|
161449 TCAACAGTATTTTCAAAATCAAGAGAAATCCGAGTTATGAGGGCCA 161498
119 sLys.....GluAsnAsnAlaLeuLeuArg.... 127
::|||
161549 TCATCATCTTCCAGTGAGGTAGTATGGCTTAGCGACCATCTTTGGTG 161598
131 AspArgAsp.....AspProSerAspAlaLeuSer 140
|::|
161599 GATAGACGCTGTCTATATACCTCGCTGAATCATCATTAATGCTATTATCA 161646

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12210
```



```

; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12214
; LENGTH: 666448
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12214
```

```
alignment_scores:
  Quality: 85.50      Length: 118
  Ratio: 1.082        Gaps: 2
  Percent Similarity: 66.949  Percent Identity: 23.729
```

alignment_block:

US-09-041-994-2_copy_613_752 x US-09-335-032-12214/rev ..

Align seg 1/1 to reverse of: US-09-335-032-12214 from: 1 to: 666448

```

18 AspaAspArgIlyHisSerSerLeuThrAsnSerProLeuAspSerSercy 34
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
613459 GAAGAGAAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 613410
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
34 slyGluSerSerValSerValThrSerProserGlyValSerSercert 51
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
613409 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 613360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
51 hrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuGln 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
613359 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 613310
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
68 GluLysHisArgIleLeuHisLysLeuGlnAsnGlyAsnSerProAl 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
613309 GACTCTGAATCATGCTCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCAG 613260
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
84 a.....GluValAlaLysIleThrAlaGlnAlaIleThrGlyL 96
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
613259 TTCACAGTGAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 613210
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
96 ysaAspThrSerSerIleThrSerCysGlyAspGlyAsnValIlyAsGln 112
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
613209 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCGAA 613160
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
113 .GluGlnLeuSerProLysLysLysGluAsnAlaLeuLeuArgTyrL 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
613159 TCAGAAATGTAACCAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGATGC 613110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
129 eu 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
613109 TA 613108
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12212
seq_documentation_block:
; Sequence 12212, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12218
```

```

; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12212
; LENGTH: 439885
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12212
```

```
alignment_scores:
  Quality: 85.00      Length: 105
  Ratio: 1.232        Gaps: 3
  Percent Similarity: 65.714  Percent Identity: 28.571
```

alignment_block:

US-09-041-994-2_copy_613_752 x US-09-335-032-12212/rev ..

Align seg 1/1 to reverse of: US-09-335-032-12212 from: 1 to: 439885

```

23 SerSerLeuThrAsnSerProLeuAspSerSercysLysGluSerSercy 39
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25662 AGCTCTCAGTCAGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 25613
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
39 l.....SerValThrSerProserGlyValSerSerThrSerclyg 54
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25612 CAGTCATCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25563
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
54 lylValSerSerThrSerAsnMetHisGlySerLeuGlnGlyLysHis 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25562 CAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25513
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
71 Arg.....IleLeuHisLysLeuGlnAsnGlyAsnSerProAl 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25512 TCAGCTCTGTGATGATCAAGCTGTGTGATGATCAATCTCTCTCTCTCTCTCTC 25463
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
84 aglValAlaLysIleThrAlaGlnAla..... 93
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25462 TGATGTCTCAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25413
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
94 .....ThrGlyLysAspThrSerSerIleThrSercys 104
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25412 CAAGCTGTGTGATGATCAAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25363
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
105 glyAspGlyAsnVal 109
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25362 GCGAGCCCAATCAGTC 25348
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:US-09-335-032-12218
seq_documentation_block:
; Sequence 12218, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12218
```



```

79130  TCTTCTCAACTGACGAACCGCTTCTCCGCAAAAAGCTCATCGATCAGACGAG  79179
32  rSerCysLysGluSerSer.....ValSerValThrSerProSerGlyu  47
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
79180  TTCACTAAGAGAGAGATAGTAAGACTAGTATTCACCCGCTTCATCATCATCAT  79229
47  aIseSerSerThrSerGlyGlyValIseSerThrSerIasMetHisGly  63
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
79230  CATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT  79279
64  SerLeuLeuGlnGlnLysHisArgIleuHisLysLeuGlnIasnGln  80
:::  |||||  |||||  |||||  |||||  |||||  |||||  |||||
79280  GAT.....AAAGTCGATCTCTCCGT  79299
80  yAsnSerProAlaGluValAlaLysIleThrAlaGlnIAsnThrGlyLysA  97
:::||||  :::  |||||  |||||  |||||  |||||  |||||  |||||
79300  ATCCCTCTTCGCGTCACTTCCCAACATCCAAACATCAATCTTCTGTAAACGGGA  79349
97  sPthSerSerIleThrSerCysGlyAspGlyAsn  108
:::|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
79350  GCGCCTCCAGCAGTACAAAGCTCAATAGAGCGGTAAT  79384

seq_name: /cgrn2_6/ptcodata/2/pma/US08_NEW_COMB.seq:US-08-956-171C-480

```

```

: seq. documentation block:
: Sequence 480, Application US/08956171C
: GENERAL INFORMATION:
: APPLICANT: Charles Kunsch
:           Gil H. Choi
:           Patrick S. Dillon
:           Craig A. Rosen
:           Steven C. Barash
:           Michael R. Fannon
: TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
: NUMBER OF SEQUENCES: 5235
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 Inch, 1.4MB storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/956,171C
: FILING DATE: 20-Oct-1997
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/009,861
: FILING DATE: January 5, 1996
: APPLICATION NUMBER: 08/781,986
: FILING DATE: January 3, 1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Hoover, Kenley K.
: REGISTRATION NUMBER: 40,302
: REFERENCE/DOCKET NUMBER: PB248P1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 610-5790
: TELEFAX: (301) 309-8439
: INFORMATION FOR SEQ ID NO: 480:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 6309 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: double
:   TOPOLOGY: linear
: SEQUENCE DESCRIPTION: SEQ ID NO: 480:
: IS-08-956-171C-480

```

[illegible]

alignment_block:
US-09-041-994-2_COPY_613_752 x US-08-956-171C-480

Align seg 1/1 to: US-08-956-171C-480 from: 1 to: 6309

```

6 H1SLYSLYSLENGLEULEUETHRCYSSERISERISPAAPATGSLYH1 22
   |||.....|.....|.....|.....|.....|.....|
82 CATAGACGATTATTGTAAAGACACCAAAATTAACGATGCTTAGGTCG 131
22 sSerSerLeuThrAnSerProLeuAspSerSerCysLysGluSerServ 39
   |||.....|.....|.....|.....|.....|.....|
132 CGGTTCAATTACTGCATTACCAATTATGGAACA...CAACCTGGTGATA 178
39 aISerValInhrSerProSerGlyValISerSerSerThrSerGlyGlyVal 55
   |||.....|.....|.....|.....|.....|.....|
179 TTTCACGCTTATGTCACACAACGCTTATTTCATTAACGATGAGACAAATC 228
55 .....|.....|.....|.....|.....|.....|.....|

```

```

229 TTCTTACATCTGATTATTATCTTCACAGTGTAAAGACACGATTAAATGC 278
56 ...SerSerThrSerAsnMetHisGlySer.....64
      :::::|||||:|||||
279 CGCACAATCTGTATCTCGTGTGGGAGATCTGCACAAATTAACCAATGA 328
65 .....LeuLeuGlnGlyHisArgIleLeu 73
      :::::|||||
329 AGAAGTTCGTGACGTTACGTCCTGTGACTTACGCGTCATACAGAACTT 378
74 HisLysLeuLeuGlnGlnAsnGlyAsnSerProAlaGluValAlaIleIle 90
      :::::|||||
379 GAATCTTTGCCAAATTGGGT.....TCAGACCTTGATGAATTAC 419
90 rAlaGlnAlaThrGlyLysAspThrSerSerIleThrSerCysGlyAspG 107
      :::::|||||
420 TCGAAGTAAATTAAGAAGGTGAACGTATCGT.....453
107 LysAsnValValLysGlnGlnGlnLeuSerProLysLysGlnAsnAsn 122
      :::::|||||
454 ..GAAGTCTTAAACAGATCAAAACAAACACATTACCGAGAACACCAA 501
124 AlaLeuLeuArgTyrLeuLeuAspArg.....AspAsp 134
      :::::|||||
502 GTGTTGATTATTATTCGATTAAACAAGATATTATAGATGAT 543

```

seq_name: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:US-08-956-171C-220

```

seq documentation block:
: Sequence 220. Application US/0895611C
:
: GENERAL INFORMATION:
:
: APPLICANT: Charles Kunsch
:
: Gil H. Choi
:
: Patrick S. Dillon
:
: Craig A. Rosen
:
: Steven C. Barash
:
: Michael R. Fannon
:
: TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
:
: NUMBER OF SEQUENCES: 5255
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: Human Genome Sciences, Inc.
:
: STREET: 9410 Key West Avenue
:
: CITY: Rockville
:
: STATE: Maryland
:
: COUNTRY: USA
:
: ZIP: 20850
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Diskette, 3.50 Inch, 1.4mb storage
:
: COMPUTER: HP Vectra 486/33
:

```



```

      |||
95162 CTTGAAGCCCGCAGCTGACGGTCTCCGATCCCGCCGACGGTTTA 95207
      |||
seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:us-09-335-032-12216
seq_documentation_block:
: Sequence 12216, Application US/09335032
: GENERAL INFORMATION:
: APPLICANT: Velculescu, Victor
: APPLICANT: Vogelstein, Bert
: APPLICANT: Kinzler, Kenneth
: TITLE OF INVENTION: Characterization of the Yeast
: TITLE OF INVENTION: Transcriptome
: FILE REFERENCE: 01107.78572
: CURRENT APPLICATION NUMBER: US/09/335.032
: CURRENT FILING DATE: 1999-06-16
: PRIOR APPLICATION NUMBER: US 60/035,917
: PRIOR FILING DATE: 1997-01-23
: PRIOR APPLICATION NUMBER: US 09/012,031
: PRIOR FILING DATE: 1998-01-22
: NUMBER OF SEQ ID NOS: 12219
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 12216
: LENGTH: 924430
: TYPE: DNA
: ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12216

alignment_scores:
      Quality: 78.00      Length: 100
      Ratio: 1.418      Gaps: 3
      Percent Similarity: 55.000      Percent Identity: 30.000

alignment_block:
US-09-041-994-2_copy_613_752 x US-09-335-032-12216/rev ..
Align seg 1/1 to reverse of: US-09-335-032-12216 from: 1 to: 924430

      23 SerSerLeuThrAsnSerProLeuAspSerSerCysLysGluSerSerGly 39
      ||||| :|||:||||| :|||:||||| :|||:|||||
879914 TCTTCAGATTCATCTCACCCTCATCTCATCTCAATTCATCTTCAGC 879865
      39 lSerValThrSerProSerGlyValSerSerSerThrSerGlyValSer 56
      ||||| :|||:||||| :|||:||||| :|||:|||||
879864 TTCATCTCCGCTTCATCTAGTATCAGCTCCGCTGCTCAAGGCTATTA 879815
      56 ex...SerThrSerAsnMetHisGlySerLeuLeuGlnGluLysHisArg 71
      :||| :||| :||| :||| :||| :|||
879814 CTTACAGTCTCTTACATGATGAGGCTCCTGT..... 879783
      72 lLeuHisLysLeuLeuGlnAsnGlyAsnSerProAlaGluValAlaArg 88
      :||| :||| :||| :||| :||| :|||
879782 .....AAATCTACTGCTCAAGTCCGCTC 879760
      88 glIleThrAlaGlnAlaThrGlyLysAspThrSerSerIle.....Ths 103
      ||||| :|||:||||| :|||:||||| :|||:|||||
879755 AGATTTAGAACAGTGTGCTGTTTATGACAACATCAGATTATATGGCGT 879710
      103 exCysGlyAspGlyAsnValLysGlnGluGlnLeuSerProLysLys 119
      :|||:||||| :|||:||||| :|||:||||| :|||:|||||
879709 ACCTGTATGACGTTGATGATGTTCTTCAAGCTTAACCTCAAGCCAGAAA 879660

seq_name: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:us-09-335-032-12219
seq_documentation_block:
: Sequence 12219, Application US/09335032
: GENERAL INFORMATION:
: APPLICANT: Velculescu, Victor
: APPLICANT: Vogelstein, Bert
: APPLICANT: Kinzler, Kenneth
: TITLE OF INVENTION: Characterization of the Yeast
: TITLE OF INVENTION: Transcriptome
```

```

: FILE REFERENCE: 01107.78572
: CURRENT APPLICATION NUMBER: US/09/335.032
: CURRENT FILING DATE: 1999-06-16
: PRIOR APPLICATION NUMBER: US 60/035,917
: PRIOR FILING DATE: 1997-01-23
: PRIOR APPLICATION NUMBER: US 09/012,031
: PRIOR FILING DATE: 1998-01-22
: NUMBER OF SEQ ID NOS: 12219
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 12219
: LENGTH: 948061
: TYPE: DNA
: ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12219

alignment_scores:
      Quality: 79.50      Length: 129
      Ratio: 1.187      Gaps: 7
      Percent Similarity: 51.938      Percent Identity: 29.457

alignment_block:
US-09-041-994-2_copy_613_752 x US-09-335-032-12219/rev ..
Align seg 1/1 to reverse of: US-09-335-032-12219 from: 1 to: 948061

      29 ProLeuAspSerSerCysLysGluSerValSerValThrSerProSe 45
      ||| :||| :||| :||| :||| :|||
295710 CCTGATGACATGATTAAGCGAGTCTCTCATTTCCGCTG..... 295672
      45 rGlyValSerSerThrSerThrSerGlyValSerSerThrSerAsnMetH 62
      |||||:|||||:|||||:|||||:|||||:|||||
295671 .GCTATTTCAGTAGACCAACCAATCGGCCCAAGTCACCAATTCACACATTA 295623
      62 lsgLysSerLeuLeuGlnGluLysHisArglLeuHisLysLeuLeuGln 78
      ||| :||| :||| :||| :||| :|||
295622 AGGACCCGTTGTACTAAGAAAGCTTAAGAAATATATCTCGCCCTGGAA 295573
      79 Asn..... 79
      |||
295572 AATCAAAAGTGGCTTTAAGGAGCGCTGTAAGCATCAGCAAGAACTCA 295523
      80 .....GlyAsnSerProAla...GluValAla.....LysIleThrA 91
      :|||:||||| :|||:||||| :|||:||||| :|||:|||||
295522 TTTTGGCTGTAATTAAGAAATAGCCCTTGAAGTGAACCAATTTTAAGACTGCTT 295473
      91 lacIlnAlaThrGlyLysAspThrSerSerIleThrSerCysGlyAspGly 107
      :|||:||||| :|||:||||| :|||:||||| :|||:|||||
295472 CGGATGACAGCGGCTTAAGCTTACTGAA.....AAGTCATCAGAAAGCC 295432
      108 AsnValValLysGln...GluGlnLeuSerPro..... 117
      :|||:||||| :|||:||||| :|||:||||| :|||:|||||
295431 ACCATTACGAAGCAAGAAAGAGAGTGTCAACCTCCCAATATGTTGAACSTGA 295382
      118 .LysLysLysGluAsnAsnAlaLeuLeuArgTyrLeu 129
      :|||:||||| :|||:||||| :|||:||||| :|||:|||||
295381 TCGGAGCTTGAGAAATTAATTAACCTTGACAGAAACCTC 295345
```


This Page Blank (uspto)